CUL-H02971-5-P000091

35 No. 1 (New Series No. 163) October 1974

ANALYSIS

Edited by
C. J. F. Williams

5

CONTENTS

Truth-tables and truth	Michael Cohen
Tendencies, frequencies and classical utilitarianism	T. S. Champlin *and* A. D. M. Walker
Identifying numbers	Gilbert Harman
Anscombe, Hume and Julius Caesar	Donald W. Livingston
Formal justice and rule-change: a reply to Holborow	T. D. Campbell
Meaningfulness without confirmability—a reply	R. G. Swinburne
Dale on material implication	Colin Roberts
Incomplete symbols again—a reply to Mr. Urmson	R. K. Perkins, Jr.
On the slippery slope again	Anne Lindsay

ASIL BLACKWELL . ALFRED STREET . OXFORD

Sixty pence

ANALYSIS 35.1 OCTOBER 1974

105
An13
VOL-35
Nos. 1-4
OCT 1974 - MAR 1975

TRUTH-TABLES AND TRUTH

By Michael Cohen

WITTGENSTEIN'S *Tractatus* 6.111 begins:

> Theories which make a proposition of logic appear to have content are always false. One might think, for example, that the words 'true' and 'false' signify two properties among others, and then it would seem a remarkable fact that every proposition possesses one of these properties.

In her *Introduction to Wittgenstein's Tractatus* Miss Anscombe is unclear about the nature of the confusion that Wittgenstein is talking about; this shows in her use of an argument put forward by Professor Geach in an article 'Ifs and Ands', an argument which betrays just the confusion about logic that Wittgenstein is attacking.

We can ask what it is that makes the interpretation of a particular piece of symbolism a part of *logic*, for the same symbolism could be understood in many ways that clearly have nothing to do with logic. Consider the following table:

p	q	pIq
1	1	1
1	0	0

In his article[1] Professor Geach suggests the following non-logical interpretation: let '1' and '0' stand respectively for the presence and absence of some hereditary characteristic in animals, and let 'pIq' mean 'offspring by p out of q'. Then the table gives us the substantive information that this characteristic is present in the offspring when the sire has it if and only if the dam has it too.

Geach is here attacking Lewis and Langford for their idea that logic is in essence a game with marks on paper. Lewis and Langford hold that while p and q might be 'any kinds of things' pIq is an 'operation or move' which can be made when p has some property A only if q has the property too. The table then tells us that if p has the property and pIq is an allowable move then q has the property. The logical interpretation then consists in taking this property to be truth.

Geach objects to this—rightly—on the grounds that it involves an inconsistency in the interpretation given to the numbers. In the columns on the left '1' and '0' stand for the presence and absence of some property (in a particular case, truth); but on the right they stand for the allowability and non-allowability of moves.

---

[1] P. T. Geach, 'Ifs and Ands', Analysis, vol. 9, no. 3, 1948–49 (reprinted in *Logic Matters* by P. T. Geach, pp. 194–198).

If we are to give a consistent interpretation to the numbers, then on any interpretation *p*, *q* and *pIq* must be the same sort of thing; but Lewis and Langford speak as if *p* and *q* could be propositions while *pIq* is not a proposition but a move. Of course if we stick to truth and falsity as the meanings of '1' and '0' then *pIq* must be not a move but a proposition—the *move* that Lewis and Langford are interested in is the move from the conjunction of *p* and *pIq* to *q*. Their difficulty is connected with a general confusion about "material implication" which I do not want to go into here.

I said that *p*, *q* and *pIq* must be the same sort of thing on any interpretation. Clearly that is not sufficient to make the interpretation a logical one since that condition is fulfilled by the interpretation in terms of hereditary characteristics. But Geach is not clear about this. He says that if *pIq* is to be a move in a game then so must *p* and *q* be, if '1' and '0' are to stand consistently for the allowability and non-allowability of moves. What the table would then show is that if a move *p* is allowable then *q* is allowable if and only if *pIq* is also allowable. And he assumes that this is sufficient to give a *logical* significance to the table. The significance is logical he says, because the sign '*I*' involves no reference to a particular game as say '*Kt*' does to chess, but only to games in general. 'It expresses what we might fittingly call a logical relation of the move *pIq* to the moves *p* and *q*; so that the significance of "*pIq*" is after all "logical" ' (*ibid.*).

In fact the interpretation is nonsense. Since a game is determined by what moves are allowed in it, it makes no sense to speak of a game in which the allowability of a move is conditional on the allowability of other moves. A move is either allowed or it is not. (Certainly the allowability of a move might be conditional on whether other moves have been in fact made in a particular game—as for example with castling in chess; but that is not at all the same thing.)

But the confusions I am concerned with go deeper than this. Miss Anscombe reproduces Professor Geach's argument and suggests that what makes the interpretation logical is that it does not give 'substantial information' about games in the way that the 'hereditary characteristic' interpretation gives information about animals. What the table does, she says, is to *define* 'a possible allowability in terms of given allowabilities, and *pIq* is a notation in which this allowability is set forth . . . '.[2] She goes on to remark that while it would be substantial information about a particular game, first that a move whose allowability was so conditioned occurred in it and second what the move actually was, still '. . . that any move describable as *pIq* is allowable, if *p* is allowable, only if *q* is also allowable, is not a "substantial" piece of information' (*ibid.*).

---

[2] G. E. M. Anscombe, *An Introduction to Wittgenstein's Tractatus,* Hutchinson University Library, 2nd edition (revised) 1963, p. 55.

But this is just the point at which the analogy between the "move" interpretation—even had this made sense—and the propositional breaks down. Given the definition of *pIq* in terms of allowability-conditions I can know what *moves p* and *q* are without knowing what move *pIq* is, or even if there is such a move in a particular game; but on the propositional interpretation of the table, given the definition of *pIq* in terms of truth-conditions, if I know what propositions *p* and *q* are I cannot go on to ask what proposition *pIq* is, or whether there is such a proposition.

The confusion behind the use of this analogy is the one Wittgenstein was talking about—that of taking truth and falsity to be properties of propositions. If truth were a property of propositions it might look as if one had to investigate, say, conjunctions of propositions to see whether they do in fact have that property if and only if their conjuncts do. That is obviously an absurd idea. But it is inadequate to say that we don't need to investigate because we have *defined* the conjunction of a set of propositions as that proposition which is true if and only if each member of the set is true; for then it is not clear why we don't need to investigate to see which proposition the conjunction of a set of propositions actually is.

This consideration makes obscure the point of Miss Anscombe's claim that the legitimacy of the truth-table explanation of a truth-function—conjunction, say—depends on the existence of some unique proposition which is the conjunction of a set of propositions; for it leaves unclear how knowing that such a unique proposition exists is connected with knowing what proposition it is. I know, for example, that for any set of numbers there exists a unique number which is their sum, even though there are infinitely many sets of numbers whose sum I do not know. Here we speak of discovering the sum of a particular set by calculation rather than investigation; but it makes no sense to speak in general of calculating the conjunction of a set of propositions.

In order to understand a non-logical interpretation of the tables I must already grasp the logic of truth-functions. (This has a certain property *if* that has the property *only if* the other *also* has the property.) But in the propositional interpretation the tables tell us no more than I must already know in order to be able to understand them. What then is the point of truth-tables?

Tractatus 4.43-4.44 says:

> We can express agreement with truth-possibilities by correlating the mark 'T' (true) with them in the schema. The absence of the mark means

> disagreement . . . the sign that results from correlating the mark 'T' with the truth-possibilities is a propositional sign.

A pair of independent propositions has four truth-possibilities:

p	q	
T	T	T
T	F	F
F	T	F
F	F	F

The possibility that p is true just is the possibility that p; the possibility that p and q is true is just the possibility that p and q, and so on. And this makes sense only on the propositional interpretation. (That a move is allowable is not itself a move.) If I now write the sign '$p \,.\, q$' above the column of T's and F's on the right I am simply introducing a piece of notation. We would not say that I am introducing a concept, for, as we have seen, this 'definition' could not be understood by someone who did not already understand the truth-functions. What is important in the notation is not the sign ' . ' but rather what is common to all the signs that express the same thing: '. . . in "$\sim p$" it is not "$\sim$" that negates; it is rather what is common to all the signs of this notation that negate p' (5.512). Truth-tables are a way of representing 'what is common'.

In saying that the possibility that p is true just is the possibility that p, I am attributing to Wittgenstein a view of truth sometimes called the redundancy theory. In his article 'Truth', and again in his recent book on Frege, Michael Dummett has argued that the redundancy theory is incompatible with the view—which Wittgenstein certainly held—that to understand a proposition is to know in what conditions it is true and false.

> . . . if *all* that it means to say that 'Frege died in 1925' is true is that Frege died in 1925, then the knowledge that 'Frege died in 1925' is true just in case Frege died in 1925 is simply the 'knowledge' that Frege died in 1925 just in case Frege died in 1925.[3]

One might as well argue that if all it means to say 'n is prime' is that n has only itself and unity as factors then the knowledge that n is prime just in case n has only itself and unity as factors is simply the "knowledge" that n has only itself and unity as factors just in case n has only itself and unity as factors.

Such an argument would show that there could be no definition of concepts at all. In fact Dummett's case for the inadequacy of the redundancy theory's 'definition' of truth (when that definition is coupled with the idea that understanding a proposition is knowing its truth-conditions) depends on analogy with a case where defining a term simply by saying in what cases the term was correctly applied is inadequate as an explanation of what the term means.

---

[3] M. Dummett, *Frege: Philosophy of Language*, Duckworth, 1973, p. 458.

> . . . if the whole explanation of the sense of the word 'win' consisted in a stipulation, for each game, of the conditions under which one player or side was said to have won, then a knowledge of what a particular game is could not involve knowing what it is to win that game . . . (*ibid.*).

The point, made paradoxically here, is that someone could know how to pick out the winner of a game just if he knew for example which positions are called 'winning positions'; but

> . . . if he was even unaware that games were competitive activities in which each player or side was striving to win, he would rightly be said to have only a partial understanding of the word 'win' . . . (*op. cit.*, pp. 459-60).

While he could, that is, apply the word to the right cases, he would have no idea of the significance of the distinction which its use marks.

The inadequacy of this definition of 'win' depends on the possibility of teaching someone to recognize a winning position without teaching him that winning is the point of the game. But this possibility does not exist for 'true'; one could not learn to apply the word 'true' without knowing what difference there was between thinking a proposition true and thinking it false, the significance of that distinction. In this sense there could be no definition of 'true proposition' as there is say of 'prime number'—nor does the redundancy theory offer one. To suppose that there could be such a definition is to suppose that one could know what a proposition is without having the concept of truth, just as one could have the concept of number without having the concept 'prime number'.

The point of the redundancy theory lies in its denial of the idea that truth and falsity are properties of propositions. Far from being incompatible with the idea that understanding a proposition is knowing in what conditions it is true, it is the only account of truth which makes this intelligible, for, as we have seen, if truth were a property of propositions one could know the truth-conditions of the proposition $p . q$ without knowing what proposition it actually was.

Truth-tables do not of course provide us with definitions of the truth-functional concepts; or, better, the signs for truth-functions do not denote concepts at all. This is the point that Wittgenstein calls the 'fundamental idea' of the *Tractatus*. Wittgenstein was dissatisfied with Russell's logical symbolism in various ways. One misleading thing about it is that the signs 'v', '$\supset$' and so on look like signs for relations. But also the variety of the symbolism hides the internal connections between the various forms. 'When we infer q from $p \vee q$ and $\sim p$, the relation between the propositional forms of "$p \vee q$" and "$\sim p$" is masked in this case by our mode of signifying' (5.1311). This makes it look as if we need rules of inference to justify the step from $p \vee q$ and $\sim p$ to q. Wittgenstein rejects this in 5.132: 'The nature of the inference can be gathered only from the two propositions . . . "Laws of inference", which are supposed to justify inferences, as in the works of Frege and Russell, have no sense and would be superflous'. If $p \, . \, q$ is explained as the proposition which is true when p is true and q is true it adds nothing to this explanation to be told that p follows from $p \, . \, q$. Truth-tables simply provide a uniform notation for this kind of definition.

The truth-table presentation of logic is often contrasted with the so-called 'axiomatic development'. But the contrast is misleading. Lewis Carroll brought out the futility of thinking of the 'axioms' of logic as premisses of inference—but I think it has not always been clear what the confusion in this is. In *The Principles of Linguistic Philosophy* Friedrich Waismann remarks that the axiomatic presentation of logic makes it look like

> a deductive theory the propositions of which are linked together by proofs; it seems in this respect to resemble exactly other deductive systems such as mechanics and geometry.[4]

But Waismann fails to explain how logic differs from mechanics and geometry. He says that 'since logic is to justify inference it must consist not of propositions but of rules' (*ibid.*). The rules, Waismann thinks, determine the senses of the symbols ' . ' '$\sim$' and so on. But we can think of the axioms of geometry in just the same way as rules of inference, thus: 'from a statement of the form "x, y are points" you may infer one of the form "There exists a straight line between x and y".'

It will be said that the rules of inference of a geometrical symbolism do not by themselves determine the sense of expressions like 'line' and 'point' but that we need to give the symbolism a certain sort of application. The same holds true in a sense for logical symbolism; we need to understand the signs like 'p' and '$p \vee q$' as standing for *propositions*. But there is an important difference between the two sorts of application, and it is a failure to appreciate this which underlies the theories of logic that Wittgenstein was attacking. In learning geometry one has to learn the application of expressions like 'straight line'; one has to learn, that is, what it is for propositions of the form 'Points x, y and z determine a straight line' to be true or false. But there is nothing like this with logic.

It might be thought that just as in learning geometry one learns when to call a line 'straight' so in learning logic one learns when to call a proposition true (one learns for example to call the proposition 'not-p' true just when p is false and false when p is true). But this is misleading, for one could know what a line is without having the concept 'straight';

---

[4] F. Waismann, *The Principles of Linguistic Philosophy*, Macmillan, 1965, p. 377.

but one could not know what a proposition is without having the concepts of truth and falsity (cf. *Tractatus* 4.063).

In proposing the idea of alternative logics philosophers have often appealed to an analogy between logic and geometry. Just as the replacement of Euclid's parallel axiom leads to the development of non-Euclidean geometries, so it is argued that the replacement (or simply the dropping) of certain axioms of standard logic will lead to the development of other logics. But it is not clear what the significance of these systems is supposed to be. Waismann says that 'an alteration in the calculus alters the rules of inference and thereby the grammatical rules which govern the use of the word "proposition" ' (*op. cit.*, p. 382). But the word 'proposition' is not a part of logical symbolism; and the idea that different logical calculi could determine different uses of the word 'proposition'—in the way that different geometries can be said to determine different uses of 'straight line'—is as obscure as the idea that the point of our ordinary logical calculus is to determine the use of the word 'proposition'. It might be said that it is more in line with our analogy to suppose that—just as the term 'line' can be thought of as having a constant meaning in different geometries, while the sense of 'straight line' changes—so there is one concept of proposition common to all logics, while different logics give different senses to 'true proposition'. But the implication of this, that just as a line might be straight from the point of view of one geometrical system and not from that of another, so one and the same proposition might be true in one sense of 'true' and not in another, is unintelligible, if to know the sense of a proposition is to know when the proposition would be true, when false.

Wittgenstein opposed any account of logic which made it appear to be a matter of general truths about such concepts as negation—as if in learning logic I were interested in learning the application of expressions like 'the negation of *p*' (as if that were the description of some proposition which bore a relation to *p*). If truth and falsity were properties of propositions, the truth-tables would not in any case determine what proposition was the negation of *p*. One might rather say that logic is supposed to tell me what it is, for example, to deny—or assert—a proposition; and that of course is something which cannot be told.

University College of North Wales, Bangor

TENDENCIES, FREQUENCIES AND CLASSICAL UTILITARIANISM

By T. S. Champlin *and* A. D. M. Walker

I

IN his recent *Utilitarian Ethics* (Macmillan, 1973) Anthony Quinton at one point invokes an argument which has long enjoyed the status of orthodoxy in utilitarian exegesis. Speaking of the objection that 'utilitarian ethics lays an impossible burden of calculation on the moral agent', he refers us to Mill's reply that

> accumulated experience provides us with knowledge of the moral *tendency* of actions of particular kinds, embodied in rules or principles that are subordinate to the principle of utility itself.

And he continues:

> This is one of two clear pieces of evidence that Mill was a rule-utilitarian ... The other place where he appears to subscribe to the rule-utilitarian doctrine is in his main formulation of the principle of utility, where he says that an act is right if it *tends* to promote happiness. An individual action cannot have a tendency. Producing certain effects more often than not cannot be a characteristic of an individual action which occurs once and once only and has one and only one set of effects. Only a kind or class of actions can have a tendency to promote happiness or anything else (*op. cit.* pp. 47-8, Quinton's italics).

This argument has, of course, a distinguished history. Couched in terms of a distinction between type- and token-actions (rather than particular actions and classes or kinds of actions), it was applied to these same passages of Mill by J. O. Urmson in 'The Interpretation of the Moral Philosophy of J. S. Mill' (*Philosophical Quarterly* 1953, and reprinted in P. R. Foot (ed.), *Theories of Ethics*; the argument occurs on p. 135 of the latter). And some ten years later Mary Warnock, in her Introduction to the Fontana edition of *Utilitarianism*, argued in a similar vein that Bentham's use of the noun 'tendency' would suggest that he too was a rule-utilitarian (*op. cit.* p. 22).

As a matter of fact we believe that the force of this argument can be undermined by a careful examination of the writings of Bentham and Mill themselves: there are passages in Bentham's *Principles* and in Mill's *Logic*, and perhaps even in *Utilitarianism* itself, where the language of 'tendency' is clearly used in connection with a particular or token-action (see, for example, *An Introduction to the Principles of Morals and Legislation*, edited by J. H. Burns and H. L. A. Hart, pp. 11-12, 74-5, 79, 127, 145; *System of Logic* III 10, 5; and *Utilitarianism*, Fontana edition, p. 268). Our purpose in this note, however, is not to engage in textual velitation but to expose and discuss a fallacious *philosophical* assumption common to the several versions of the argument.

II

The orthodox argument, as we shall call it, can be reduced to the following propositions:

(i) An understanding of what is meant by statements of the form '. . . tend(s) to φ' or '. . . have (has) a tendency to φ' shows that the analysis of such statements must make reference to some class or kind.

(ii) Thus when a 'tendency'-statement has a grammatically singular subject, as in 'A tends to φ', the subject-term must be interpreted as referring to the class of A's, not to a particular A, to the type A rather than to a token A.

(iii) Hence when utilitarians speak of the tendency of an action to, or of an action's tending to, promote happiness, they must have in mind a type or kind of action.

The fallaciousness of the inference from (i) to (ii), which is perhaps already evident enough, can best be brought out by considering the philosophical account of tendency which appears to be at the root of the orthodox argument. This is the view according to which "tendency"-statements are to be analysed in terms of "*frequencies*", statements of the form 'A's tend to be B's', for example, being analysable as, roughly, 'Most A's are B's' or 'A's are usually (more often than not) B's' (see K. R. Popper, in S. Körner (ed.), *Observation and Interpretation*, pp. 65-70; D. H. Mellor, *The Matter of Chance*, p. 69; A. J. Ayer, *Probability and Evidence*, especially pp. 61-3). To be sure, this view of tendency is not without its difficulties, and has been attacked by, among others, P. T. Geach (see G. E. M. Anscombe and P. T. Geach, *Three Philosophers*, pp. 103 ff). But for present purposes we shall simply ignore the question of its acceptability. Our primary concern is not with the truth of the frequency account but with the fallacious inferences which have been drawn from it by Quinton and his predecessors.

Now admittedly the move from (i) to (ii) does have a certain appeal. Indeed if one concentrates on such an example as 'Red squirrels tend to live in North America' it is easy to be tempted into concluding that either 'tend' must take a plural subject or that when it takes a singular this must refer to a type rather than a token (as in 'The red squirrel tends to live in North America'). But that this conclusion, despite its attractiveness, is unwarranted can immediately be seen from the possibility of such statements as 'Smith tends to wear a red tie' or 'Jones tends to eat porridge for breakfast'. These can be glossed, in conformity with the 'frequency' account, as 'Smith usually wears a red tie' and 'Jones more often than not eats porridge for breakfast', but there is no doubt that in each sentence the subject-term refers to a particular individual. And the explanation—which is at the same time the reason for the fallaciousness of the orthodox argument—is not far to seek: we need only recognize that the "frequencies" licensing the use of 'tend' may be introduced on (so to speak) either the subject or the predicate side of a sentence. That is to say, the class within which the required "frequencies" are to be found may be either the class denoted by the subject-term of the original sentence or some class associated with the predicate—for instance, the class denoted by the *object* of the sentence (as in 'Robinson tends to have irresponsible tenants'). To produce counter-examples it is necessary, therefore, merely to introduce the "frequencies", as we have done, on the predicate side.

A defender of the orthodox argument may reply that the demonstrated flexibility of the "frequency" account does not, after all, appreciably diminish the strength of his position. For he may concede the possibility of "tendency"-statements in which the subject refers to a particular, but still deny the existence of examples where the subject refers to a particular *action*. Indeed, he will add, the absence of such examples is precisely what the "frequency" account would lead one to expect: it is precisely because Smith as an individual persists through time and so has many opportunities of wearing a tie that he can be described as tending to wear a red one; but a particular action, not similarly persisting, cannot have many occasions of φ-ing or not φ-ing. The defence, therefore, consists in offering a revised formulation of the argument: although the "frequency" account does not justify the *general* claim that "tendency"-statements cannot have a subject referring to a particular, it does, when fortified by further premisses about the peculiar nature of actions, justify the *specific* claim that a "tendency"-statement cannot have as subject a particular *action*.

But this defence suffers from an obvious defect. The impermanence of a particular action, by contrast with the temporal durability of many particular objects, is not sufficient to prevent the former from figuring as the subject of a "tendency"-statement. Paradoxical as it may sound, with some values of φ, it is no barrier to X's φ-ing on many different occasions that X does not exist on these occasions. Thus a politician's cowardly action may repeatedly attract the criticism of his compatriots, or a man's filling in a form incorrectly may on numerous subsequent occasions lead to inconvenience. This last example should give pause to the interpreter of Bentham and Mill: for if a particular action may produce effects at different times over a long period, this provides a natural context in which a utilitarian may wish to make remarks about the tendency of a particular or token-action. Be this as it may, the fact remains that an action's not persisting through time is not a satisfactory ground for arguing that particular actions, unlike particular objects, cannot be the subject of "tendency"-statements.

III

In conclusion we may deal briefly with two responses to our claims in the last section, the responses, as we may characterize them, of the blunt sceptic and the sophisticated sceptic.

The strategy of the blunt sceptic is to concede the weakness of the orthodox argument but to doubt the value of our refutation. The unsoundness of an argument, he will doggedly insist, must not be equated with the falsity of its conclusion; and is it not significant that no counter-examples have as yet been produced to destroy decisively the specific claim on which Urmson, Mary Warnock and Quinton must rely? But this direct challenge can be as directly met. Would there be any unintelligibility in a man's claiming that suddenly to abolish surtax next Monday would tend to promote euphoria on the Stock Exchange; that it might also tend to push up share prices, or to harden the attitude of the Trade Unions? Again, I can wonder whether my offering to resign would tend to improve matters, or I may be confident that what my friend did last Friday has definitely tended to make things worse. And surely a bystander may say that for me to interfere at this stage might only tend to annoy my squabbling neighbours even more. The abolishing, offering, doing and interfering mentioned in these statements are not kinds or classes of actions; they are individual, datable actions, actually performed or possibly to be performed; they bear the imprint of the particularity of the token, not the generality of the type.

The reaction of the sophisticated sceptic is rather different. Appealing to the doctrine that the grammatical subject of a sentence cannot always be identified with its logical subject, he will take exception on this ground to at least some of the counter-examples we have offered in the preceding pages. Thus with 'Jones tends to eat porridge for breakfast' he may argue that the proper logical form is 'Jones's breakfasts tend to be porridge-eating occasions', and in this we no longer have as subject a particular but a class. The straightforward reply to this must be that, whatever the *truth* of the sophisticated sceptic's claim—and the whole question of logical subjects seems to us exceedingly problematic—his manoeuvre cannot help because it constitutes an *ignoratio elenchi*. The orthodox argument is, after all, intended to resolve the interpretation of certain texts where an action-word occurs as the *grammatical*, not the *logical* subject. Granted that in these texts, as well as in other examples, the action may not be the *logical* subject, this cannot be of the least assistance to a philosopher who holds that these texts exclude an act-utilitarian interpretation, for what he must maintain is that the *grammatical* subject cannot be a particular or token-action. *His* argument is, therefore, conclusively destroyed by showing, as we have done, not only that the "frequency" account can allow for "tendency"-statements with a particular action as grammatical subject, but that there is in fact no difficulty in supplying concrete examples of such statements.

University of Hull

IDENTIFYING NUMBERS

By Gilbert Harman

To say that numbers can be identified with sets is not, as Schumm and Stevenson suggest (Analysis 33.3, commenting on my 'A Non-essential Property', *Journal of Philosophy* 67 (1970): 183–185), to say either that numbers *are* sets (which is too strong) or that number theory has a model in set theory (which is too weak). It is to say, rather, that any ω sequence of different sets (or other things) can be *used* as the numbers. With this in mind, we might consider the following hypothesis concerning the logical grammar of talk about numbers: numerals are best analysed as function symbols rather than names.

Knowing that $5+7=12$, we know that for any such sequence s, the sum operation for that sequence applied to the fifth and seventh members of the sequence yields the twelfth member.* We can make this more explicit by writing

$$5_s +_s 7_s = 12_s.$$

Here '5', '7' and '12' are not names but function symbols. The letter 's' names a sequence and '5_s' names the fifth member of that sequence; but we need not suppose that '5' by itself names anything. It is the 'f' in 'fx'. In particular, to say that numerals are really function symbols would not be to say that they are names of functions. They would not be names at all. It follows that they would not be names of objects (which confirms Benacerraf's conclusion in 'What Numbers Could Not Be', *Philosophical Review* 74 (1965): 47–73). It also follows that they would not be names of objects that have essential properties.

---

* The sum operation '$+_s$' can be recursively defined in the usual way in terms of more basic operations. If $N_s(x)$ is the successor of x in the sequence s (of distinct things) and 1_s is the first member of s, $x+_s 1_s=N_s(x)$ and $x+_s (y+_s 1_s)=N(x+_s y)$.

Princeton University

ANSCOMBE, HUME AND JULIUS CAESAR

By DONALD W. LIVINGSTON

RECENTLY,[1] Professor Anscombe has argued that Hume's account (in Section IV, Part III, Book I of the *Treatise*) of why we believe that Caesar was killed in the senate house on the ides of March is incredible since it entails that we must treat the existence of Caesar as an hypothesis that can be freely doubted. But we can throw so fundamental a belief into 'question only by indulging in Cartesian doubt' (*ibid.*, p. 6) which would entail 'destroying bases and standards for discovering any historical facts at all' (*ibid.*, p. 7). In what follows, I shall argue that Hume's analysis of historical method does not commit him to this absurdity. It is worthwile to point this out because Hume was one of the first philosophers to do what has come to be known in the last decade or so as "critical" or "analytical" philosophy of history. Yet the literature on Hume's philosophical reflections on history is virtually non-existent. It is for this reason that Professor Anscombe's article is especially welcome, and this is also why it is important to get Hume's views straight.

I

Two passages are in question. The first is about how we believe in matters falling outside our experience and memory:

> When we infer effects from causes, we must establish the existence of these causes . . . either by an immediate perception of our memory or senses, or by an inference from other causes; which causes again we must ascertain in the same manner, either by a present impression, or by an inference from *their* causes, and so on, until we arrive at some object which we see or remember. 'Tis impossible for us to carry on our inferences *in infinitum*; and the only thing, that can stop them, is an impression of the memory or senses, beyond which there is no room for doubt or enquiry. (Selby Bigge's edition, pp. 82–3.)

As an application of this analysis Hume has us consider

> any point of history, and consider for what reason we either believe or reject it. Thus we believe that CAESAR was kill'd in the senate-house on the *ides of March*; and that because this fact is establish'd on the unanimous testimony of historians, who agree to assign this precise time and place to that event. Here are certain characters and letters . . . the signs of certain ideas; and these ideas were either in the minds of such as were immediately present at that action . . . or they were deriv'd from the testimony of others, and that again from another testimony . . . 'till we arrive at . . . eye witnesses and spectators of the event. 'Tis obvious all this chain of argument or connexion of causes and effects, is at first founded on those characters or letters, which are seen or remember'd. (Ibid., p. 83.)

---

[1] G. E. M. Anscombe, 'Hume and Julius Caesar', ANALYSIS 34.1, October, 1973.

Anscombe's interpretation runs as follows: (1) the second passage is about an inference from an effect (an historical record) to its remote cause (Caesar's death) and so is not an application of the principle stated in the first passage which is about inferring effects from causes. But Hume's slip here can be easily amended since for him 'cause and effect are inferentially symmetrical' (Anscombe, p. 2). (2) Hume's claim in the first passage: ''Tis impossible for us to carry on our inferences *in infinitum*' Anscombe takes to mean '*the justification of the grounds of our inferences cannot go on in infinitum*' (Anscombe's italics, p. 2). (3) Hume is, therefore, requiring that there be a starting point in the justification of our inferences which is either perceived or remembered. And it is natural to say that in history the starting point is a perception or memory of a contemporary record. But this is not Anscombe's interpretation:

> Hume is arguing not merely that we must have a starting point, but that we must *reach* a starting point in the justification of these inferences. He would have been clearer if he had said, not 'we cannot carry our inferences on *in infinitum*' but 'we cannot trace them back *in infinitum*'. But . . . cause and effect . . . are inferentially symmetrical. So for him the tracing back is inference too. But note that it must be purely *hypothetical* inference. (Ibid., p. 2.)

Hume's method then is to begin with 'the mere idea of Caesar's death' (*ibid.*, p. 3), and then reason hypothetically that if Caesar were killed there would have been eyewitnesses, and if so there would have been testimony from the eyewitnesses, and if so there would have been records, and if so . . . until we reach a consequent proposition about the record in hand which, having perceptual content, we can assert. (Had there been no perceived record our chain of inference would have gone on *in infinitum*.) We now begin with a true antecedent (the proposition about what the record says) and work back to the proposition about Caesar's death which we can now assert. This is indeed an incredible procedure, and Anscombe rightly points out that 'Belief in recorded history is on the whole a belief *that there has been* a chain of tradition of reports and records going back to contemporary knowledge; it is not a belief in the historical facts by an inference that passes through the links of such a chain' (*ibid.*, p. 4).

But Hume is not committed to this absurdity. Notice that in the second passage the reason for believing the proposition about Caesar's death is not that it is the result of a long chain of causal reasoning but simply that it is 'establish'd on the unanimous testimony of historians, who agree to assign this precise time and place to that event'. The statements which immediately follow about the chain of testimony and records constitute not a justification of belief but an analysis of the structure of historical beliefs. One must keep in mind the limited purpose of this section of the *Treatise* which is entitled 'Of the component parts of our reasonings concerning cause and effect', and which is only a page and a half long. The main point is sketched out in the first sentence: 'Tho' the mind in its reasonings from causes or effects carries its view beyond those objects, which it sees or remembers, it must never lose sight of them entirely, nor reason merely upon its own ideas, without some mixture of impressions, or at least of ideas of the memory' (*Treatise*, p. 82). The contrast here is between pure *a priori* reasoning and reasoning that is disciplined by experience of the world, the main point of the section being that 'the component parts' of causal reasoning (the cause or effect) must at some point make contact with our experience of the world. But there is another important point Hume wants to make: one who reasons causally *believes* he is reasoning about real existence. Now according to Hume's mechanical theory of belief (which he hopes will both explain the existence of causal beliefs and provide norms for regulating them) all beliefs about real existence are caused ultimately by impressions. (We need not examine here Hume's complicated defence of this thesis nor the numerous qualifications he makes to it.) Yet it is not necessary to have either the original impression or the memory of it to reason causally. As he says in the concluding paragraph: 'For even supposing these impressions shou'd be entirely effac'd from the memory, the conviction they produc'd may still remain' (*ibid.*, p. 84). What Hume is anxious to show here is that causal reasoning begins and ends in a context of *conviction* about real existence and that we are aware of this when we reason causally. On the other hand, we are aware that we do not believe the results of '*hypothetical* arguments . . . there being in them, neither any present impression, nor belief of a real existence' (*ibid.*, p. 83).

So when Hume says we cannot carry our inferences out *in infinitum* he does not mean as Anscombe interprets him that the justification of our grounds cannot continue *in infinitum* and, consequently, must have a starting point (much less that we must first *reach* a starting point by hypothetical inference); rather, he means (a) that if we reason hypothetically (*a priori*) it will be arbitrary to stop at one point rather than another, and (b) that wherever we stop we cannot (given Hume's theory of belief) believe what we say: 'Every link . . . wou'd in that case hang upon another . . . and consequently there wou'd be no belief nor evidence' (*ibid.*).

Applying this interpretation to the second passage, Hume's view is that history is based on records (written or oral) which are perceivable and so public. Without records historical reasoning would be *a priori* and could not terminate at a non-arbitrary and believable point. But mere inscriptions or sounds in the present are not enough for historical knowledge: they must be interpreted. And the interpretation Hume places on them is that they express 'the unanimous testimony of historians' about the facts of Caesar's death, i.e., they are understood to express propositions within the body of what we believe to be historical knowledge. But this means that the inference Hume is talking about is directly from the document to the proposition about Caesar's death and not as Anscombe supposes an inference that passes through the links of a causal chain. Oddly, this is very like the view Anscombe wishes to defend in the passage quoted above and the following:

> We know Caesar from the testimony of ancient historians, we even have his own writings! And how do you know *that* these are ancient historians . . . ? You were told it . . . . We know it from being taught; not just explicit teaching, but by its being implicit in a lot else that we are taught explicitly. But it is very difficult to characterize the peculiar solidity involved, or its limits. (Anscombe, pp. 5–6.)

Even so. And like everyone else, Hume has his difficulties explicating this 'peculiar solidity', but he does affirm it. This can be made more clear by considering Hume's remarks on history in sections of Part III other than the Section IV we have been examining.

II

(1) In Section VII, Hume brings up again the argument of Section IV that without a perceptual content causal reasoning would go on *in infinitum*, making it clear that what he has in mind is what was argued for above, namely, the distinction between empirical and *a priori* reasoning:

> When we infer the existence of an object from that of others, some object must always be present either to the memory or senses . . . since the mind cannot run up with its inferences *in infinitum*. Reason can never satisfy us that the existence of any one object does ever imply that of another. (*Treatise*, p. 97.)

As an instance of this, Hume again turns to a reflection on history. He points out that there is no intrinsic difference between the ideas of a novel and those of a history. The difference is extrinsic and consists in the fact that we believe the historian but give 'no credit to the testimony of the author' of the novel. And we believe the historian because we accept him as part of our tradition of knowledge about the past. Hume, of course, was aware that it must be possible to criticize the tradition. But it is beyond the scope of this paper to examine how he thought this could be done, except to call to mind, in passing, his famous argument in the *Enquiry Concerning Human Understanding* that laws of nature may be used as canons to exclude as historical evidence all testimony about miracles.

(2) In Section IX Hume makes a special point to argue that, although a testimony or record is an effect of a cause which it is the testimony or

record of, it is a very special effect: 'Other effects only point out their causes in an oblique manner; but the testimony of men does it *directly*, and is to be consider'd as an image as well as an effect' (*ibid.*, p. 113, italics mine). This passage should be compared with Anscombe's challenge to what she thinks is the Humean analysis of historical knowledge: '*If* the written records that we now see are grounds of our belief, they are first and foremost grounds for belief in Caesar's killing .... Then belief in that original event is a ground for belief in much of the intermediate transmission' (Anscombe, p. 3). But this is just Hume's view of the matter: the record in hand of Caesar's death is an image through which we see 'directly' not a long string of documents but Caesar's death.

(3) In Section XIII, Hume again returns to our belief in the existence of Julius Caesar, this time raising an imaginary sceptical objection against his own theory of belief. The objection is that since any record about a 'point of ancient history' must have passed through 'an almost immeasurable length' of oral and written testimony and copies of editions, it would seem that no matter how strong the vivacity of our original belief in the record in hand, it must diminish in strength as we reflect on the 'many millions' of transfers which have to be truth-preserving if the record in hand is to be believed. Consequently, it would appear that all belief in ancient history 'will be lost in time, as the chain of causes encreases, and runs on to a greater length'. But Hume holds it 'contrary to common sense to think, that if the republic of letters . . . continue . . . our posterity, even after a thousand ages, can ever doubt if there has been such a man as *Julius Caesar*' (*Treatise*, p. 145). Hume's solution is derived from his 'image' analysis of historical records mentioned above:

> tho' the links are innumerable, that connect any original fact with the present impression, which is the foundation of belief; yet they are all of the same kind, and depend on the fidelity of Printers and Copists. One edition passes into another, and . . . so on, till we come to that volume we peruse at present. There is no variation in the steps. After we know one, we know all of them; and after we have made one, we can have no scruple as to the rest. (Ibid., p. 146.)

Whatever difficulties there may be in this account, it is clear that, for Hume, our belief in the existence of Caesar is grounded in a fundamental belief in the historians who have written the book we are reading and not in a conclusion reached by inference through the links in a chain of record. In the light of this belief, the complicated connections between the original event and the book in hand collapse, rendering the latter, in effect, a first hand report.

But if an historical document is to be thought of primarily as an "image" through which we directly see the original past event, why does Hume talk about 'a chain of arguments of almost an immeasurable length' connecting the present record with the original fact? It is this sort of talk that makes Anscombe's interpretation appear plausible. The main reason is Hume's conception of what it means to have a conception of (or more precisely to *imagine*) a past existent.

> 'Tis also remarkable, that in the conception of those objects, which we regard as real and existent, we take them in their proper order and situation, and never leap from one object to another, which is distant from it, without running over, at least in a cursory manner, all those objects, which are interpos'd betwixt them. When we reflect, therefore, on any object distant from ourselves, we are oblig'd not only to reach it at first by passing thro' all the intermediate space betwixt ourselves and the object, but also to renew our progress every moment; being every moment recall'd to the consideration of ourselves and our present situation. (Ibid., p. 428.)

On this view, the idea of Caesar is a temporally complex idea entailing the idea of our present situation and the idea of the correct temporal order between Caesar and ourselves. When Hume speaks of 'passing thro' all the intermediate space betwixt ourselves and the object', he does not mean that we must methodically *infer* our way through it as a condition of our knowing the object but, rather, that the idea of the intermediate space is entailed in the idea of the object. Hume makes the same point about memory: 'The chief exercise of the memory is not to preserve the simple ideas, but their order and position' (*ibid.*, p. 9). But, for Hume, there is an important difference between the idea of Caesar and an idea of memory. The latter is an idea of (a) something that happened to us along with (b) the primitive force and vivacity of the idea through which it is believed. The idea of Caesar is the idea of (a′) a past existent which we can neither see nor remember, associated in the imagination with (b′) the idea of a tradition of sound testimony and records through which it is believed. So anyone who thinks of Caesar's existence historically must also actually *imagine* 'at least in a cursory manner' the tradition of sound testimony and records through which the idea of Caesar is believed. Hume grants that we cannot have a very clear idea of the tradition, but that is no matter: since the links 'are perfectly resembling, the mind runs easily along them, jumps from one part to another with facility, and forms but a confus'd and general notion of each link' (*ibid.*, p. 146).

On Hume's view, then, we cannot separate in imagination the idea of Caesar from the tradition of testimony and record through which he is known without ceasing to have the idea of Caesar that we have. So Hume is not committed to treating Caesar as an hypothesis we have framed to explain later events as Anscombe suggests in her challenge to what she thinks is Hume's account: 'People "in history", as we say, are not in any case hypotheses which we have arrived at to explain certain phenomena' (Anscombe, p. 7). Such views of history are common in contemporary analytical philosophy of history, a recent example being Professor Murphy's thesis that 'George Washington enjoys at present the epistemological status of an electron: each is an entity postulated for the purpose of giving coherence to our present experience, and each is unobservable by us'.[2] Whether true or not, this pragmatic and instrumental account of historical thinking should be quarantined as far as possible from interpretations of the Humean perspective.

---

[2] Murray G. Murphy, *Our Knowledge of the Historical Past* (Bobbs-Merrill: New York, 1973), p. 16.

Northern Illinois University

FORMAL JUSTICE AND RULE-CHANGE: A REPLY TO HOLBOROW

By T. D. CAMPBELL

IN ANALYSIS 34.2 Les Holborow argues against my thesis that rule-change involves formal injustice by attacking my rejection of one type of alleged exception to this thesis, namely 'those cases where the existing rule explicitly contains a time clause, e.g. if rule 1 states that only group A, being persons in category C before *t*, come within its scope, then group B, being persons in category C after *t* are specifically excluded from the rule and it is not therefore formally unjust to treat them in accordance with rule 2' ('Formal justice and rule-change', ANALYSIS 33.4, p. 114). Against this alleged exception to the formal injustice of rule-change I argued that:

> Where a particular time is mentioned in a rule as, for instance, where a time limit of a specific date is incorporated within it so that the rule is restricted to a named historical period, then formal injustice arises because of the lack of generality in the rule. From the point of view of the rule of law, rules containing a specific time reference are not genuinely general rules, and to apply such defective rules is tantamount to formal injustice since the distinction between group A and group B is arbitrary as opposed to rule governed. (p. 114.)

Holborow cites as a counter-example 'the case of those emergent African states which required European and Asian colonists to decide by a certain date whether they wished to become citizens of the new state or alternatively to retain their British citizenship' (p. 61). In such a case, Holborow argues, the reference to a specific time is neither irrelevant nor arbitrary, and this counter-example shows that provisions cannot 'be shown to be formally unjust merely on the ground that they contain a particular time-reference' (p. 61). I shall argue that Holborow's objection does not touch the substance of my thesis and requires, at the most, a minor modification in its presentation.

Assuming, for the moment, that cases of the type outlined by Holborow are, as he assumes, not formally unjust, the absence of formal injustice here might be explained by the fact that the specific time mentioned in the law in question marks a change in the situation of the persons who are affected by the existing rules. As I wrote in my original paper 'it should be noted that the thesis does not state that it is formally unjust to introduce a new rule to meet a new situation' and the emergence of a new state in a colonial territory may be thought to constitute just such a new situation. As Holborow points out, 'the provision is concerned with a once-and-for-all event, the establishment of a new state', and this is an event which seems to alter the situation of the pre-independence colonists in such a way as to make it inappropriate to say that they are in the same essential category as those post-independence colonists who remain in the new state; formal injustice is thereby avoided. Moreover the fact that the relevance of the reference to a specific date in the citizenship provision is due to the change in circumstances brought about by the creation of the new state demonstrates that the example does nothing to show that the passage of time is in itself a factor which can exempt a change in rules of this type from formal injustice. So, while it is not strictly correct to say that a rule which contains a specific time reference thereby always involves formal injustice, it is accurate to say that the mere fact that a rule does contain a specific time clause limiting the period of its application is not in itself sufficient to make the changes in treatment consequent upon these limitations compatible with formal justice.

However it may be doubted whether the change in situation brought about by the emergence of a new state in a colonial territory is relevant to the situation of pre- and post-independence colonists in such a way as to avoid the formal injustice of such new citizenship laws. If the colonists are simply regarded as residents within the territory in question then the mere fact of a constitutional change in the sovereign authority may be regarded as insufficient to justify making a distinction between "colonial" and "non-colonial" residents. Thus from the point of view of the law of that territory as it existed prior to independence (and presumably as it in very large part remained in force after independence) persons in the same essential category (residents) are being treated differently owing to the introduction of a new law and this is formally unjust. Or, to put it in another way, the post-independence colonists might justifiably regard it as unfair that they should have to make a choice of citizenship not required of pre-independence colonists. Of course there may be many good moral reasons favouring the creation of a new state which override the consideration that this will involve formal injustice to some residents within the pre-independence territory, but the formal injustice of the rule changes following from the granting of independence may still be regretted.

If this is the correct analysis of such a situation then the reference to a specific time in the citizenship provision may be regarded, not so much as the cause of formal injustice, but as an attempt to minimize the harm done by the formal injustice which is consequent on the rule-changes which are part and parcel of the creation of a new state. As I argued in my original paper 'the formal injustice of rule-change provides grounds for implementing such changes in a gradual manner so as to minimize the contrast in the treatment accorded to groups A and B' (p. 118). In this instance the citizenship provision, by allowing for a choice of citizenship on the part of colonial residents and by giving them a reasonable time within which to make their choice, could be regarded as an attempt to reduce the hardships arising for colonists out of the formal injustice of the rule-changes resulting from independence. In this case the example cited by Holborow as an exception to my thesis may be taken as providing some support for that thesis since it provides an illustration of an actual attempt to minimize the suffering caused by the formal injustice of a certain type of rule-change.

University of Stirling

MEANINGFULNESS WITHOUT CONFIRMABILITY —A REPLY

By R. G. SWINBURNE

IN the course of 'Confirmability and Factual Meaningfulness' (ANALYSIS, vol. 33.3 (1972–3), pp. 71–6) I argued that the confirmationist principle is false. This is the principle that a statement is factually meaningful if and only if either it is itself an observation statement or there are observation statements which, if true, would confirm or disconfirm it. By 'confirm' I understand 'raise the probability of', and by 'disconfirm' I understand 'lower the probability of'. I urged that conclusive counter-examples to the principle can be produced. These are propositions which state that some claims about the unobserved having such-and-such probability on the best evidence which men will ever obtain, have such-and-such a truth-value (e.g. are true). Such propositions, I claimed, are factually meaningful (given that their referring expressions succeed in referring), but neither confirmable nor disconfirmable. As my main example I discussed

> p_1: Among possible claims about the prehuman past which the best evidence ever to be obtained by man makes highly improbable some are nevertheless true.

R. I. Sikora (in 'Confirmability and Meaningfulness', ANALYSIS, vol. 34.4 (1973–4), pp. 142–4) has attempted a rebuttal of my arguments to show p_1 neither confirmable nor disconfirmable. (I shall normally in future abbreviate 'neither confirmable nor disconfirmable' by 'unconfirmable'.) Sikora claims that my arguments to show p_1 unconfirmable are inadequate, and that in fact p_1 is disconfirmable. I claim that Sikora's main argument to show p_1 disconfirmable fails, because he has misunderstood a simple point in my argument, one which may have been expressed unclearly. I acknowledge however that one consideration which Sikora brings forward necessitates a minor amendment to my example. I go on to provide for the amended example a more rigorous proof of its unconfirmability.

In support of the unconfirmability of p_1, I claimed that to have evidence for p_1, you would have to have evidence about some claim q showing it to be a claim which the best evidence ever to be obtained makes highly improbable, evidence which increased the probability of its truth while showing that description to continue to apply to it. I argued (p. 75)—no doubt in too loose a way—that 'there cannot be any evidence which counts for a claim q being a claim which the best evidence ever to be obtained makes highly improbable and at the same time counts for q being true', and I claimed that 'from this it follows that no evidence can count for p_1'. Sikora misunderstands this claim when he represents me as claiming that 'you would have to show that a statement having the following form could be made more or less probable:

> Claim q (a claim about the prehuman past) is highly improbable in terms of the best evidence ever to be obtained by man, and claim q is true.'

Sikora, that is, represents me as claiming that in order to (dis)confirm p_1 we would need to (dis)confirm a conjunctive statement of the above form, statements of which form he calls Q-statements. He represents me (p. 142) as claiming that 'no evidence can count for or against the conjunction as a whole'. He then goes on to suggest that we could disconfirm a Q-statement by disconfirming one of its conjuncts without necessarily confirming the other to an equal extent (or conversely). But I was not claiming that evidence could not count for or against Q-statements, and although I admit that my argument was too loose, I do not think that it is naturally susceptible of Sikora's interpretation. p_1 concerned claims about the prehuman past which the best evidence ever to be obtained by man makes highly improbable. To confirm p_1 you would need to get evidence first that a claim q is of this character and then evidence that q is true; and to disconfirm it you would need evidence that a claim q had that character and then evidence that q is false. If evidence shows that q has not the former character, it shows that it is not a claim of the kind p_1 is concerned with. You need evidence about what the best evidence will show about q before evidence about q's truth is shown relevant. But, my argument went, necessarily if your present evidence shows q to be a claim very improbable on the best evidence, it thereby also gives q a constant degree of improbability on present evidence— and so while the evidence shows q to be a claim of the kind we are concerned with, q is unconfirmable. If evidence turns up raising or lowering the probability of q, it thereby counts against q being a claim of the kind with which p_1 is concerned. That is why, I argued, p_1 is unconfirmable.

In the course of his argument (pp. 142, sq.) Sikora suggests that difficulties arise for my example owing to the fact that the term 'highly improbable' covers a range of probability values. These I overlooked. Sikora's suggestion can in fact be developed to produce a powerful argument against the use of my example to falsify the confirmationist principle. For, given Sikora's point, one could obtain evidence which slightly raised the probability of some claim q while leaving q highly improbable. Since the evidence left q highly improbable (and since 'highly improbable' covers a range) it could leave it as probable as before that q was 'a claim highly improbable on the best evidence which men will ever obtain'. Yet the evidence confirms q. Hence the evidence confirms the suggestion that a claim highly improbable on the best evidence which men will ever obtain is true, and so confirms p_1. I overlooked this difficulty because I wrongly thought of 'highly improbable' (and the similar expressions used in my other examples) as denoting an exact probability value.

In view of this difficulty I need to amend my example so that the probability expression in it denotes an exact probability value (other than 1 and 0). The expression need not denote a probability value to which a number can be assigned. We do not normally give exact numerical probability values (e.g. '0.01') to historical claims, and it would in general seem highly arbitrary to do so. One can however assign exact probability values other than numerical ones. Instead of 'makes highly improbable' I could write 'makes as probable as is the claim on present evdence that a rocket left Cape Kennedy for the moon in 500 m.B.C.' or 'makes as probable as is on present evidence the claim that dinosaurs never lived on Earth'. Further however there is one exact numerical value which we can fairly naturally assign to historical claims—'0.5', i.e. 'as probable as not'. Some claims are as likely as their negations. My amended thesis is then that any claim of the form p is unconfirmable:

> p: among possible claims about the prehuman past to which the best evidence ever to be obtained by man gives probability x some are true.

where x is a particular exact probability value other than 1 or 0. I will take 'x' as '0.5' or 'as probable as not'; but if anyone wishes to fill out 'x' with a different exact probability value he may do so—he will need to make one or two obvious consequential amendments to arguments below.[1] I will call a claim about the prehuman past to which the best evidence ever to be obtained by man gives a probability of x an x-claim. My amended thesis is that p, the proposition that some x-claims are true, while factually meaningful (given that any referring expressions contained in them succeed in referring) is neither confirmable nor disconfirmable. Given that reference succeeds, how could it fail to be factually meaningful to assert that some of the claims about the prehuman past with a certain probability on certain evidence are true? (Even if there is no prehuman past or there are no such claims about it, it still makes sense to suppose that there are these things, and so to suppose that reference succeeds.)

In the original article I supported my assertion of the unconfirmability (of my original p_1) by claiming (very loosely) that since necessarily any evidence which made it probable that a claim was 'a claim which the best evidence ever to be obtained by man renders highly improbable' would render it probable that it was false, we could not obtain evidence that a claim of the former kind was true. Sikora has rightly pointed out that my argument was too loose. I welcome the opportunity provided by his article to raise the level of sophistication of the debate, and to produce, for my amended example, a more rigorous proof of its unconfirmability.

---

[1] A claim is in general to be understood simply as a statement; but in the case where 'x' is taken as '0.5' a claim must be understood in a more restricted sense, unless p is to become an analytic truth. For the negation of any statement about the prehuman past which has a probability of 0.5 will also be a statement about the prehuman past with a probability of 0.5. Since either a statement or its negation must be true, if a claim is taken as any statement and 'x' as '0.5', p will be analytic and not factual. The obvious restriction is to take a claim as any existential assertion, any statement asserting the existence of something (e.g., 'Once upon a time there were dinosaurs on Earth'). Then the negation of a claim about the prehuman past will not itself be a claim about the prehuman past, and p will remain factually meaningful.

p claims that some x-claims are true. It thus has the form 'some A's are B'. To confirm such a proposition involves adding to the probability of some object which is A (whether a named object or any such object) being B; that is it, involves passing from having evidence e to having evidence e^1 where, for some z, $P(Bz|Az.e^1) > P(Bz|Az.e)$. You can only confirm 'some cows are white' if you can get evidence which makes it more likely than before that some object which is a cow is white; and if you cannot get the latter evidence you cannot confirm 'some cows are white'. (Confirming 'some cows are white' is to be contrasted with the very different problem of confirming 'some objects are white cows', i.e. 'there are white cows'.) So to confirm p one would have to be able to get evidence which added to the probability of some claim which was an x-claim being true. If one cannot do that one cannot confirm p. Let q be a claim about the prehuman past, and e be our present evidence relevant to its truth. We may suppose that its probability on e has some value y—$P(q|e)=y$. Now to confirm p one would have for some such claim q to pass from evidence e to evidence e^1 such that $P(q|e^1.q$ is an x-claim$)$ $>P(q|e.q$ is an x-claim$)$. That however one could never do. For any claim q about the prehuman past and for any evidence e or e^1 which we could have now or in future, $P(q|e^1.q$ is an x-claim$)=P(q|e.q$ is an x-claim$)=x$. For if we know what the best evidence will show about some claim q, the present probability of q must be the same as on that evidence —whatever else we know now relevant to the truth or falsity of q. If we know, that is have conclusive evidence, that when detectives have found out everything ascertainable their evidence will make it probable that Jones is guilty, then our present evidence makes it probable that Jones is guilty. If we know the results of future inquiries, they form part of our evidence, and if we know that they will be based on the best evidence which men will ever have, then we can conclude that their conclusions are as probable as that future evidence will show them to be. So however e may vary, $P(q|e.q$ is an x-claim$)$ continues to equal x. Hence the probability of q, on the supposition that it is an x-claim, remains constant and so q is unconfirmable on that supposition. Hence p also is unconfirmable. (True, if either e or e^1 were to entail either q or $\sim q$, then the above probabilities would not have the value x. But e and e^1 being evidence which we could have now or in future, neither e nor e^1 will entail

q or $\sim q$—since e and e^1 are our present or future evidence and q and $\sim q$ claims about the prehuman past.)

It might seem to an opponent that we could obtain evidence for or against some *x*-claims being true. Taking x as 0.5, we might study what subsequent evidence came to show about past claims which had at the time a probability of 0.5. We might find that most claims with a probability at one stage of 0.5 turned out to be false. We might discover this in the course of a more general discovery that most reasonable beliefs which men had held in the past turned out to be false. This evidence of past human fallibility is indeed relevant to assessing future claims, but the import of this evidence is not what an opponent supposes it to be. It does not show that future claims with a probability on the total evidence available in future of 0.5 are less likely to be true. For this new evidence of human fallibility becomes part of the evidence which we will have for assessing future claims. And that means that claims which without the new evidence would have a probability of 0.5 will be less likely to be true than that, because the total evidence (including the evidence of human fallibility) will give them a probability less than 0.5. The new evidence means only that different claims will have a probability of 0.5 on the total evidence, different ones from those which would otherwise have this probability. What the evidence could not show is that claims which will have on the best evidence (including the evidence of human fallibility) available in future a probability of 0.5 are more likely or less likely than that to be true—for it could only do that if the new evidence included better evidence than the best.

An analogy should make this clear. Suppose that using evidence of their ancestry and their successes in earlier races we calculate the probabilities of various horses winning in a certain race at Kempton. We then find out that horses to which such calculations have in the past given a high probability of winning have never won at Kempton. This new piece of evidence does not show that it is probable that a horse which has on the total evidence a low probability of winning will win the race in question. The difference which it does make—when it is added to the previous evidence—is that a horse which on the previous evidence had a high probability of winning our race no longer has such a high probability. For the new evidence has shown that ancestry and past success don't make the difference at Kempton which we would ordinarily rightly suppose that they do.

I conclude that the basic thesis of my earlier article stands. This was that propositions which state that some claims about the unobserved with such-and-such probability on the best evidence which men will ever obtain have such-and-such a truth-value, while factually meaningful, cannot be confirmed or disconfirmed—so long as 'such-and-such probability' is understood as denoting an exact probability value. As my main example of claims about the unobserved I took here and in the earlier article claims about the prehuman past. But all sorts of other examples are possible. I gave two other examples of claims about the unobserved, p_2 and p_3, in the earlier article—these examples remain examples of my thesis—so long as the terms in them 'highly probable' and 'very probable' are understood as denoting exact values.

University of Keele

DALE ON MATERIAL IMPLICATION

By Colin Roberts

A. J. DALE, in 'A Defence of Material Implication' (Analysis, 34.3), argues that there is a difficulty in the non-truthfunctional interpretation of 'if, then'. His strategy is to define what he calls a sufficiency condition for the falsity of *if p then q* (namely that if p is true and q false then *if p then q* is false) and to claim that this is the only uncontroversial part of the non-truthfunctional interpretation. He then shows that the sufficiency condition establishes the validity of some but not all of the inferences that are intuitively valid for 'if, then', and suggests that we establish the rest by accepting the validity of the schema of exportation

If (if both A and B then C) then (if A then if B then C)

which is, he says, true to our everyday usage. The difficulty is that the schema of exportation, together with the sufficiency condition, makes 'if, then' wholly truthfunctional.

I shall say very little about the schema of exportation. Dale is correct in arguing that its invalidity is a necessary condition of a non-truthfunctional interpretation of 'if, then', and where his arguments for its validity are independent of his remarks on the sufficiency condition they have some force. My concern is to show that the insufficiency of the sufficiency condition presents no problem.

I shall use $p \rightarrow q$, $p \supset q$, $p \vdash q$ to abbreviate, respectively, *if p then q*, *not both p and not q*, *p entails q*. Dale shows that of the following pairs of intuitively valid entailments, only the right-hand one is established by the sufficiency condition, while the unestablished left-hand one is the logical law named

(Transitivity) $((p \rightarrow q) \& (q \rightarrow r)) \vdash (p \rightarrow r)$ $\qquad ((p \rightarrow q) \& (q \rightarrow r) \& p) \vdash r$
(Contraposition) $(p \rightarrow q) \vdash (-q \rightarrow -p)$ $\qquad ((p \rightarrow q) \& -q) \vdash -p$

Dale notes that the sufficiency condition will be insufficient whenever the entailed formula is of the form $A \rightarrow B$ and the entailment is not a substitution instance of one where the entailed formula is not of the form $A \rightarrow B$. So the schema of exportation is introduced to validate the inference from the schema corresponding to an entailment on the right-hand side to the schema corresponding to an entailment on the left-hand side.

But there is a perfectly satisfactory explanation of the insufficiency of the sufficiency condition in these cases. The sufficiency condition, as an examination of the definition shows, is merely a necessary condition of the truth of $p \rightarrow q$, and need not be expected to establish all schemata corresponding to valid inferences. It says, in effect, that $(p \rightarrow q) \vdash (p \supset q)$. If it were possible to establish that $(p \supset q) \vdash (p \rightarrow q)$, this would guarantee the equivalence of '$\rightarrow$' and '$\supset$' and thus prove '$\rightarrow$' truthfunctional. The schema of exportation does this, since it has as a consequence the principle that from $((p \supset q) \& p) \vdash q$ we can infer that $(p \supset q) \vdash (p \rightarrow q)$. On the evidence of the sufficiency condition alone, however, we cannot tell whether the law of transitivity is true for '$\rightarrow$', since it is possible that $p \rightarrow q$ says more than $p \supset q$, and the residue need not be transitive. If, for instance, '$p \rightarrow q$' means '$p \supset q$ and the probability of q given p is greater than 0.99', the residue is *not* transitive.

If non-truthfunctionalists want '$\rightarrow$' to be transitive and contraposible, they must use axioms. Even with these three axioms (the sufficiency condition and the laws of transitivity and contraposition) it is still possible to identify '$\rightarrow$' either with '$\supset$' or with '$\vdash$'. To distinguish '$\rightarrow$' from '$\supset$' the residue needs to be made explicit, and there have of course been many attempts to do this. But Dale seems to argue that since there is no agreement on the residue the residue cannot be a significant part of the meaning of '$\rightarrow$', and this clearly will not do.

University of East Anglia

INCOMPLETE SYMBOLS AGAIN —A REPLY TO MR. URMSON

By R. K. PERKINS, JR

MR. J. O. URMSON—in his article 'Russell's Incomplete Symbols' (ANALYSIS 33.3)—agrees with the main thrust of an earlier article of mine (ANALYSIS 32.6) to the effect that on Russell's official account, to show that 'X' is an incomplete symbol is *not* tantamount to showing that there are no Xs. But Urmson insists that Russell's use of the expression 'logical fiction' involves 'ontological implications' and that he (Urmson) is 'still willing to say that when Russell said that "X" was an incomplete symbol [or that X was a logical fiction] his point was almost invariably to throw doubt on the reality of Xs'. I am prepared to admit a certain degree of truth in what Urmson says, although the story is somewhat complex, and the nature of the ontological implications is not quite, I think, what Urmson supposes.

A fundamental point that must be kept in mind is that Russell's programme of logical construction is revisionary, i.e. it is designed to replace "illegitimate", pre-analytic notions by more "legitimate" ones. What Russell does, in effect, is to doubt—on grounds independent of, and antecedent to, a new analysis—the existence of, or the legitimacy of our belief in, Xs as thought of in some pre-analytic way. This is true of numbers, material objects and classes, to take only three well-known examples. Numbers, thought of pre-analytically, had generated a host of puzzles [2, 68–72]; material objects, thought of in a pre-analytic way (i.e. before his 1914 analysis in *Our Knowledge of the External World*), had involved a number of illegitimate assumptions, especially that of an unknowable cause of sense-data; and classes, thought of as things or individuals (as Russell had thought of them in 1903 in *Principles of Mathematics*), had encountered certain difficulties, including those connected with the puzzle of the One and the Many and Cantor's proof that 2^n is always greater than n [4, 72fn; 2, 80–1]. Generally, Russell's technique of logical construction has the effect of purging 'X' of its old, "illegitimate" meaning by treating 'X' as an incomplete symbol and redefining 'X'—or, rather, sentences in which 'X' occurs—in terms of more "legitimate" notions. Thus, talk about numbers is to be reconstrued as talk about certain kinds of classes; talk about material objects is to be reconstrued as talk about certain series of classes of sense-data; and talk about classes is to be reconstrued as talk about certain properties formally equivalent to properties determining those classes.

Given the revisionary character of Russell's programme of logical construction, it follows, I think, that if 'X' is an incomplete symbol, the question 'Are there really Xs?' is (*pace* Mr. Urmson) ambiguous: it may mean (A) Are there really Xs as thought of pre-analytically; or, it may mean (B) Are there really Xs as thought of post-analytically? Now it's important to keep these questions distinct since Russell gives rather different answers to them.

As regards question (A), Russell's answer is officially simple and straightforward: We don't know; there may or may not be; but, in either case, we can get on without assuming (or denying) that there are such things. It's probably true, however, that some of the difficulties attending various pre-analytic notions inclined him to deviate from his official agnosticism. This seems to be especially true in the case of classes inasmuch as he undoubtedly took certain independent considerations as militating against their existence [see 2, 80–1; 3, 260; 4, 72fn].

Question (B) is more interesting and less open to a simple answer. It is, in fact, Xs in the post-analytic sense which Russell calls 'logical fictions'. Curiously enough, in the case of classes, i.e. classes as thought of in his post-analytic way (which, of course, includes material objects and numbers), Russell says both that they exist [3, 265] and that they do not exist [3, 268]. But the contradiction is only apparent as his own discussion makes clear.

When Russell says that classes do not exist he means at least two things: (a) that classes are not among individuals; and (b) that class-symbols, i.e. symbols which putatively designate classes, do not occur in the expanded (primitive) notation of *Principia.* Indeed, when sentences containing class-symbols are re-written in expanded form there will be no symbol, nor any complex of symbols, which purports to designate a class.

Point (a) is actually connected with Russell's famous theory of types according to which class-symbols (and their associated propositional functions) are of a different logical type from individual symbols, so that one cannot say 'There are classes' in the same sense of 'there are' as one can say 'There are individuals' [3, 265]. Point (b) is tantamount to the assertion that class-symbols are not logically proper names in Russell's strict sense. Both points, taken together, express, I think, what Russell has in mind when he says that classes are not part of the 'ultimate furniture of the world' [1, 182], or, as he alternatively puts it, are 'logical fictions'.

In some places, however, in calling an X a logical fiction he seems to have in mind only point (b). This is true of the passage from *Introduction to Mathematical Philosophy* [1, 182] where Russell seems to say that the putative designata of *all* incomplete symbols, including descriptions, are logical fictions. In this sense of 'logical fiction' even descriptions of sense-data would have to count their descripta as logical fictions even though such "fictions" would surely be among individuals. It was, in fact, this use of 'logical fiction' which led me to say the things which have struck

Urmson as having an unduly 'ontologically aseptic character'. I confess, however, that this sense of 'logical fiction' does not seem to be Russell's usual one.

Nevertheless, I want to insist that neither sense of 'logical fiction' (i.e. neither the sense connected with (a) and (b) taken together, nor the sense connected with (b) only) militates against the perfectly good sense in which Russell allows that classes may be said to exist. In this sense, to say that classes exist means that there are true existential formulae of the form '(∃#)(. . # . .)', where the position of '#' is occupied by a class-symbol. Indeed, in *Principia* Russell gives proofs for such formulae [4, 195], although these formulae, it is true, are not in expanded form.

In conclusion I wish to emphasize two points. First, Russell's analysis of classes after 1903 does allow that classes exist in the sense just described, although, on this analysis, classes are not among individuals; nor are they part of the 'ultimate furniture of the world'. Now, I agree with Urmson that in this regard Russell's analysis of incomplete symbols and his notion of logical fiction involve 'ontological implications'. But—and this is the second and more important point—it was not the analysis *per se* which tended to make Russell doubt that there were such things as classes as he had thought of them in *Principles*. Rather, it was a set of independent and antecedent considerations—including the puzzle of the One and the Many and Cantor's proof—which precipitated that doubt and convinced him that a reconstruction of the old notion of class was desirable.

[1] Russell, B., *Introduction to Mathematical Philosophy* (London: Allen & Unwin, 1919).

[2] Id., *My Philosophical Development* (New York: Simon and Schuster, 1959).

[3] Id., 'The Philosophy of Logical Atomism', reprinted in R. C. Marsh (ed.) *Logic and Knowledge* (New York: Macmillan, 1956), 178–281.

[4] Id. and Whitehead, A. N., *Principia Mathematica* (Paperback Edition to 56; Cambridge: University Press, 1962).

Hesser College, New Hampshire

ON THE SLIPPERY SLOPE AGAIN

By ANNE LINDSAY

IS abortion morally permissible? Joel Rudinow (ANALYSIS, 34.5) apparently thinks it is. He criticizes what he calls 'the slippery slope argument' (SSA), an argument which he regards as meant to defend an 'extremely conservative' limitation on a woman's right to control her fertility. He further says that his criticism of SSA functions as part of a 'modest defence' of abortion. I find Rudinow's comments deficient in the following two ways.

1. Rudinow mischaracterizes SSA. Borrowing from Rudinow for my own purposes, I suggest that an adequate characterization of SSA depends upon understanding what SSA is meant to establish; it is not meant to establish that human life begins at conception. Rather, attention to the contexts in which SSA is invoked discloses that SSA is meant as a *criticism* of a position which requires that we recognize a distinction between intra-uterine beings. The force of SSA as a criticism lies in exposing the lack of any basis for such a distinction. Thus, much of Rudinow's parody is beside the point.

2. Rudinow's proposed strategy in defence of abortion seems anything but modest. Rudinow will grant, I think, that a *prima facie* sufficient reason for restricting a woman's right to obtain a non-therapeutic abortion would be a demonstration that the exercise of that right involves the violation of the rights of another individual. Isn't it rather important then that we *determine* whether or not, owing to the status of the fetus, the wilful interruption of pregnancy interferes with the fetus's rights? What are we to make of Rudinow's liberal, who would sanction abortion provided only the apparent lack of an ironclad argument establishing the humanity of the fetus? The assumption underlying his strategy is that the burden of proof in the matter rests with the anti-abortionist. Rudinow's liberal is prepared to stop arguing having merely criticized one of his opponent's arguments. I think a pro-abortionist should expect to run somewhat harder than this. What of the possibility of alternative anti-abortion arguments? And what if Rudinow's criticism of SSA is wrong? If Rudinow's criticism of SSA is successful at all, then it succeeds only in showing that SSA is inconclusive as regards the status of the fetus, or that the status of the fetus is still controversial. But the very fact that the status of the fetus is controversial ought to count for something in our moral deliberations concerning abortion. Specifically, it seems morally questionable to condone a practice of killing something whose status is still in controversy. The immodesty of Rudinow's proposed defence of abortion then consists in discounting the controversy, which his own argument underscores, over the status of the fetus.

Fanshawe College

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Margaret A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, A. C. MacIntyre, Bernard Mayo, D. H. Mellor, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

SUBSCRIPTIONS. The subscription to ANALYSIS is £2.50 ($7.50) for individuals in Great Britain, £3.00 ($9.00) for individuals abroad, and £3.50 ($10.50) for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford, or placed with any bookseller.

CONTRIBUTIONS. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgment is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in ANALYSIS remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

26 OCT 1975

PRINTED IN GREAT BRITAIN BY BURGESS AND SON (ABINGDON) LTD.

ol. 35 No. 2 (New Series No. 164) December 197

ANALYSIS

Edited by
C. J. F. Williams

CONTENTS

Knowledge, ignorance and presupposition — John A. Barker
What Magellan's voyage didn't prove or why the earth is flat — Jerry S. Clegg
The examiner examined — B. H. Slater
Whiteley on rules of language — C. R. Carr
'Is morality a system of hypothetical imperatives?' a reply to Mr. Holmes — Philippa Foot
On Anscombe's exposition of Hume — A. C. Genova
Wright on functions — Patrick Grim

BASIL BLACKWELL . ALFRED STREET . OXFORD

Sixty pence

ANALYSIS 35. 2 DECEMBER 1974

KNOWLEDGE, IGNORANCE AND PRESUPPOSITION

By John A. Barker

1. SUPPOSE that Smith sincerely asserts

(1) I do not know that Brown is dead.

Is Smith's statement self-certifying?[1] Let us construe this question as: Given that Smith knows the language, can we be sure that Smith's statement is true? It may seem at first blush that we could be sure of its truth. Perhaps knowing does not entail knowing that one knows. But it seems plausible that if one knows one does not believe that one does not know. Now given the conditions described above concerning Smith's sincerity and command of the language, it appears that we could conclude that he believes that he does not know that Brown is dead. Hence we seem to have good grounds for saying that Smith does not know that Brown is dead.

2. Yet it seems odd that Smith did not instead say

(2) I do not know whether Brown is dead,

a statement which may indeed be self-certifying. Or, more exactly, assuming that Smith's statement neither contained stressed parts nor occurred in a context giving the effect of emphasized parts, it seems odd that Smith did not either use 'whether' instead of 'that', or else make one of the following stress-containing statements

(3) I do not *know* that Brown is dead
(4) *I* do not know that Brown is dead
(5) I do *not* know that Brown is dead.

(Of course, if Smith believed that he knew that Brown was *not* dead, he could not sincerely assert (2), (3), (4) or (5); but under these conditions it would nevertheless be odd for him to assert (1).)

I will argue that Smith's actual statement not only is not self-certifying, but is self-stultifying, owing to the fact that it carries conflicting implications. My strategy will be to show that this is a consequence of the truth of the general thesis that an assertion that someone knows that a certain proposition is true involves the *presupposition* that the proposition is true. I will attempt to clarify and to argue for this thesis, subsequently applying it to the case at hand.

---

[1] This was the topic of a problem symposium held at the 1973 meeting of the American Philosophical Association—Eastern Division. A shorter version of the present paper was presented at the symposium.

3. Suppose that Jones made the statement

(6) Mary does not know that Brown is dead

with no stress or emphasis involved. We can say that Jones *implied* that Brown is in fact dead, in the sense that, owing to the characteristics of the language, if Jones did not believe that Brown is dead, he used the language in a misleading way. In this sense of 'imply', which is not, of course, the logician's sense of the term, what a speaker *implies* comprises more than what a speaker would be said to *assert*.[1] But this sense of 'imply' does not extend to all propositions which could, given the speaker's statement, be legitimately taken to be believed by him. For example, suppose that the man who had won a certain lottery said

(7) The winner of the lottery is an unhappy man.

Suppose also that, while his audience believed that he was aware of the fact that he had won, it so happened that he believed that someone else had won. It is probably the case that, if the winner did not believe that he himself was an unhappy man, then, in saying what he said, he misled his audience about his beliefs. But he did not *use the language in a misleading way,* as he would have done had he said, e.g.,

(8) I am an unhappy man.

It should also be noted that in the sense of 'imply' defined above, what a speaker implies is dependent only on the characteristics of the language, and not on the speaker's intentions. For example, even if the winner of the lottery, in saying what he said, had intentionally misled his audience about his beliefs, he would not have done so by using the language in a misleading way.

4. Can we say that Smith, in asserting (1), implied that Brown is dead, just as Jones, in asserting (6), implied that Brown is dead? It may appear that there is no need to say that Smith did imply that Brown is dead, since we could, adapting a suggestion made by Jaakko Hintikka,[2] simply hypothesize that statements of the form '*a* does not know that *p*' are *sometimes* to be interpreted as meaning '*p* and *a* does not know that *p*'. But such a view cannot account for the fact that if someone were to make either of the statements

(9) If Mary does not know that Brown is dead, then Mary will attend the meeting

---

[1] Cf., e.g., P. F. Strawson, *Introduction to Logical Theory* (London: Methuen, 1952), pp.173, ff., and H. P. Grice, 'The Causal Theory of Perception', *Supplementary Proceedings of the Aristotelian Society,* XXXV (1961), 121–152, Section III.

[2] See Jaakko Hintikka, *Knowledge and Belief* (Ithaca, N.Y.: Cornell Univ. Press, 1962), pp. 12 ff., 80 ff.

(10) If Mary attends the meeting, then Mary does not know that Brown is dead

he would imply that Brown is dead. For if this person were to make either of the statements

(9′) If Brown is dead and Mary does not know that Brown is dead, then Mary will attend the meeting
(10′) If Mary attends the meeting, then Brown is dead and Mary does not know that Brown is dead

he would *not* imply that Brown is dead. Thus, the hypothesis concerning variable meaning of statements of the form '*a* does not know that *p*' cannot explain all of the relevant facts.

5. We should note that if someone made any of the following statements

(11) Mary knows that Brown is dead
(12) If Mary knows that Brown is dead, then Mary will attend the meeting
(13) If Mary attends the meeting, then Mary knows that Brown is dead

he would imply that Brown is dead. In view of those observations we should consider employing Peter Geach's notion of a 'double-barrelled assertion'—the notion of the making of a *pair of assertions,* as distinct from the making of a single assertion of a conjunctive proposition.[1] According to Geach,

> an assertoric sentence of the form '*a* has pointed out that *p*' is exponible as the double-barrelled assertion of '*a* has maintained that *p*' and of '*p*' itself. Again (an example of Frege's), '*a* fancies that *p*' is exponible as the double-barrelled assertion of '*a* thinks that *p*' and of 'it is not the case that *p*'. Assertions thus exponible will certainly retain part of their assertoric force when put, for example, into an 'if' clause; thus, one who asserts, 'If *a* is under the illusion that *p*, then *q*' does not mean 'If *a* is under the impression that *p* but it is not the case that *p*, then *q*'—rather, he both asserts 'It is not the case that *p*' and asserts 'If *a* is under the impression that *p*, then *q*'. (P. 456.)

Let us then posit as a working hypothesis that any statement containing the phrase '. . . knows that *p*' will be exponible as a pair of assertions, one of which is the assertion that *p*.

Geach, however, warns us against putting any stock in such a hypothesis: 'Use of the expression ". . . knows that *p*" does not commit the speaker to asserting "*p*"; to adapt an example of Hintikka's, one who asserted in 1916, "If Russell knows that Wittgenstein is dead, then Wittgenstein is dead" would not himself be asserting "Wittgenstein is dead"' (p. 456). Nevertheless, as D. R. Altmann has pointed out,[1] Geach's example involving 'knows that' is not parallel to his examples involving 'pointed out that' and 'fancies that'. His example involving 'knows that' is of the form 'If *a* knows that *p*, then *p*', while his other examples are of the form 'If *a* pointed out that *p*, then *q*' and 'If *a* fancies that *p*, then *q*', where '*q*' represents some proposition distinct from that represented by '*p*'. By appropriately modifying the examples, Altmann turns the tables on Geach: '"If Russell knows that Wittgenstein is dead, then I will ring him" does carry a commitment to asserting "Wittgenstein is dead", while "If Russell fancies that Wittgenstein is dead, then Wittgenstein isn't dead" does not commit one to asserting "Wittgenstein isn't dead"' (p. 175).

---

[1] P. T. Geach, 'Assertion', *Philosophical Review,* LXXIV (1965), 449–465, pp. 453, ff.

Altmann goes on to say: 'In general, even though "X" might be such that "$X \supset q$" commits one to asserting "p" . . . , "$X \supset p$" will not' (p. 175). Following this generalization we could successfully predict that someone who made the statement:

(14) If Mary knows that Brown is dead, then Brown is dead

would, in contrast with the case of (12), fail to imply that Brown is dead. But Altmann's generalization certainly does not go to the heart of the matter, as can be seen from considering the fact that a speaker making any of the following statements would likewise fail to imply that Brown is dead

(15) If Brown is dead and Mary knows that Brown is dead, then Mary will attend the meeting
(16) If Mary attends the meeting, then Brown is dead and Mary knows that Brown is dead
(17) If Brown is dead, then Mary knows that Brown is dead.

Moreover, Altmann's generalization would not cover statements (9′) and (10′) above. Indeed, the general phenomenon we are dealing with has recently been discussed at length by Lauri Karttunen and several other linguists, who have shown that it is so complex that correct generalizations about it are quite difficult to formulate.[2]

Instead of continuing our present approach and attempting to improve on Altmann's generalization, however, let us pause to consider two difficulties—the first relatively easy to handle, the second indicating that a different approach is called for. The first difficulty concerns the phenomena themselves. Karttunen's perception of the relevant facts is not fully in agreement with Geach's and Altmann's. Karttunen in effect holds that a statement such as

---

[1] D. R. Altmann, 'Geach on "Know" in "If" Clauses', ANALYSIS 33.5 (1973), 174–175.

[2] See Lauri Karttunen, 'Presuppositions of Compound Sentences', *Linguistic Inquiry*, IV, 169–193, and the bibliography appended thereto.

(14) If Mary knows that Brown is dead, then Brown is dead

involves the implication that Brown is dead. Concerning an example similar to (14), Karttunen remarks: '[The statement] is a somewhat peculiar example. There would be no point in uttering such trivial tautologies unless one were engaged in some painstaking deductive reasoning, trying to track down all the logical consequences of [the statement]' (p. 177). This remark points toward a possible explanation of the disagreement—in many (if not in most) cases in ordinary conversational contexts, a logically true 'if' statement has a force somewhat similar to that of a 'since' statement, in that the context indicates that the speaker believes that both the antecedent and the consequent are true. Evidently Karttunen failed to note that in other contexts such logically true statements have no such force. Thus, it appears that a statement such as (14) does not in itself involve an implication that Brown is dead; hence, Geach and Altmann's perceptions of the relevant phenomena seem correct.

The second difficulty concerns two inherent weaknesses in our present approach. It seems highly misleading, at best, to suggest that someone who makes the statement

(6) Mary does not know that Brown is dead

has made a pair of *assertions*, one to the effect that Brown is dead, the other to the effect that Mary does not "believe" this. It seems that the speaker at most *implies* that Brown is dead, while *asserting* only that Mary does not "believe" this.[1] A second, closely related weakness consists in the fact that, in relation to compounds of (6), the notion of 'double-barrelled assertion' does not seem to lend itself to highlighting the important difference between the two "assertions". The "assertion' that Brown is dead, unlike the assertion that Mary does not "believe" this, seems to be involved as an implication in most compounds of (6). The term 'presupposition' has often been employed by theorists in relation to certain commitments which appear to be distinct from assertions. Let us then drop the notion of 'double-barrelled assertion', and, taking a cue from a paper by Jerry L. Morgan,[2] hypothesize that statements containing the phrase '. . . knows that *p*' involve a *presupposition* that *p*.

---

[1] Cf., e.g., P. F. Strawson, 'A Reply to Mr. Sellars', *Philosophical Review*, LXIII (1954), 216–231, pp. 217, ff.; Altmann, op. cit., p. 174. I borrow from Altmann the use of 'believe' in scare-quotes explained by him in n. 1 on p. 175.

[2] Jerry L. Morgan, 'On the Treatment of Presupposition in Transformational Grammar', in R. Binnick *et al.*, edd., *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society* (Chicago: University of Chicago, 1969), 167–177.

6. To obtain any genuine philosophical benefit from this hypothesis, we must face the problem of clarifying the notoriously troublesome concept or concepts of presupposition. In its ordinary uses, the term 'presuppose' seems to have at least two main senses: (i) an 'ontological' sense of *involving as a necessary antecedent condition,* as in 'Government presupposes order'; (ii) a 'psychological' sense of *taking something for granted,* as in 'In using this ruler he presupposed its accuracy'. In discussing statements, theorists of language have formulated two more-or-less parallel conceptions: (i′) a "semantic" conception of *involving as a necessary condition of being either true or false,*[1] and (ii′) a "pragmatic" conception of *signalling something believed to be true and believed to be already accepted by one's audience.*[2] For purposes of trying to make sense out of the phenomena we have been discussing, however, each of these two conceptions has certain weaknesses.

The "semantic" conception seems illuminating if we focus narrowly of such pairs of statements as

(11) Mary knows that Brown is dead
(6) Mary does not know that Brown is dead,

both of which involve the implication that Brown is dead. If we assume that these two statements are logical contradictories—they cannot both be true nor both false—then it seems that, since Brown's being dead could be said to be in some sense a necessary condition of the truth of (6) as well as of (11), it is therefore a necessary condition of the truth and of the falsity of (11). We are then tempted to view the relation between Brown's being dead and (11) as consisting essentially in the circumstance that Brown's being dead is in some sense a necessary condition for (11) to have a truth value. From our present perspective, however, there are two problems with this "semantic" conception.

First, it in effect accords a special status to the negative logical compound of (11), namely (6). But, as we have seen, Brown's being dead is just as much an implication of some other logical compounds of (11), such as

(13) If Mary attends the meeting, then Mary knows that Brown is dead.

Why should negative logical compounds figure in the very conception or definition of the relationship in question? If we give such negative compounds a special status, will we not compound the problems of explaining and predicting the "projection" of the presuppositions of component statements on to the other logical compounds? Viewed in terms of such a conception, the patterns by which various logical compounds acquire or fail to acquire the presuppositions of their components have appeared to investigators as rather bewildering in complexity.[1] For example, why do statements (9), (10), (12) and (13) share the presupposition of (11), while statements (9′), (10′), (14), (15), (16), and (17) fail to share this presupposition?

---

[1] See, e.g., P. F. Strawson, *Introduction to Logical Theory,* pp. 175, ff.

[2] See, e.g., Robert C. Stalnaker 'Pragmatics', *Synthese,* XXII (1970), 272–289, and 'Pragmatic Presuppositions', read at the Texas Linguistics and Philosophy Conference, 1973.

The second problem with the "semantic" conception involves the assumption that (11) and (6) are logical contradictories. Intuitively, (6) seems to be too strong a statement to be a logical contradictory of (11), since someone who wished to deny (11) could make the weaker statement

(18) If Brown is dead, then Mary does not know that Brown is dead.

Since it appears that (11) is false if (18) is true, and since Brown's being dead does not seem to be a necessary condition of the truth of (18), it appears that Brown's being dead is not a necessary condition of the falsity of (11).

The "pragmatic" conception of presupposition seems illuminating if we focus on such statements as

(19) The woman who attended the meeting is now dead.

In making this statement a speaker normally signals not only a belief that the woman he referred to attended the meeting, but also a belief that his audience is already aware of this. (The signalling in question is, of course, a function of the characteristics of the language itself, rather than of the intentions of the particular speaker, or of his actually possessing the relevant beliefs.) It is the signalling of the presence of this second belief which gives to the proposition that the woman attended the meeting a kind of "background" status, a status in view of which we are led to the correct expectation that the proposition would carry over as an implication of logical compounds of (19), such as

(20) If the bomb exploded, then the woman who attended the meeting is now dead.

In contrast, the speaker of statement (19), while signalling that he believes that the woman referred to is now dead, does not signal that he believes that his audience is already aware of this. Hence, the proposition that the woman is now dead has a "foreground" status, a fact which leads us to the correct expectation that the proposition would not be an implication of such logical compounds as (20).

But if we widen our perspective and consider such statements as

(21) The woman, who attended the meeting, is now dead

---

[1] See Karttunen, op. cit., pp. 176, ff.

we find that there is normally no signalling that the speaker believes his audience is already aware of the implied proposition, namely, that the woman referred to attended the meeting. And yet the proposition that the woman attended the meeting seems to be just as much a presupposition of (21) as of (19). For example, the statement

(22) If the bomb exploded, then the woman, who attended the meeting, is now dead

clearly involves the implication that the woman attended the meeting.[1]

Thus, for our present purposes, both the "semantic" conception of presupposition and the "pragmatic" conception seem to involve difficulties which militate against their use.

7. In partial analogy with Geach's notion of 'double-barrelled assertion', let us introduce the following conception of presupposition: *a presupposition of a statement is an implication which is independent of the logical compounding operations of natural language.* If we think of the logical compounds as being built sequentially out of more elemental statements, there appear to be certain implications involved in the elemental statements which are not affected or cancelled by the natural language logical operations by which negative compounds, conditional compounds, disjunctive compounds, etc., are constructed from the elemental statements. It is precisely these implied propositions which, thereby continuing as implications of the compounds, constitute the presuppositions of the relevant statements. Thus, the selective effect of the logical compounding operations of natural language differentiates a statement's primary or focus information-content from its secondary or presuppositional information-content.[2]

In terms of this new conception of presupposition, which we should perhaps call a "syntactic" conception, the "projection" problem for presuppositions of logical compounds takes on a different appearance. Since all logical compounds will by definition share the presuppositions of their components, cases of apparent logical compounds which do not possess the presuppositions of their components should turn out on examination not to be genuine logical compounds. For example, the statements

(15) If Brown is dead and Mary knows that Brown is dead, then Mary will attend the meeting
(23) If Mary does not know that Brown is dead, then Brown is dead
(18) If Brown is dead, then Mary does not know that Brown is dead

---

[1] Cf. Karttunen, op. cit., p. 170, note 2.

[2] A similar view of presuppositions is sketched in my 'Presupposition and Entailment', forthcoming in *Notre Dame Journal of Formal Logic.*

appear to be logical compounds containing the statement

(11) Mary knows that Brown is dead

as a component, despite the fact that they all fail to presuppose what (11) presupposes, namely, that Brown is dead. But if we apply to these statements the fundamental logical principles of commutation, contraposition and exportation, we obtain

(15′) If Mary knows that Brown is dead, then if Mary does not attend the meeting Brown is not dead
(23′) If Brown is not dead, then Mary knows that Brown is dead
(18′) If Mary knows that Brown is dead, then Brown is not dead.

Thus application of these fundamental logical principles seems to be not merely invalid (i.e. not truth-preserving), but productive of some kind of nonsensical statement. Consequently, we have good evidence that the relevant compounds are not genuine logical compounds.

8. It is not difficult to find additional evidence that the type of 'and' compound found in (15), a compound we could term a *presuppositional conjunction,* is not a genuine logical compound. For example, consider the statement

(16) If Mary attends the meeting, then Brown is dead and Mary knows that Brown is dead.

If we were to treat the presuppositional conjunction contained in (16) as a genuine logical conjunction, then, by employing the logical principles of distribution and simplification, we could deduce the statement

(13) If Mary attends the meeting, then Mary knows that Brown is dead,

a statement which, unlike (16), involves the presupposition that Brown is dead. It seems clear that an entailed statement cannot possess presuppositions not involved in the entailing statement.

The following observations provide further evidence that the *presuppositional conditionals* found in (23) and (18) (and in the closely related statements (14) and (17)) are not genuine logical compounds. The oddity of the following arguments indicates that the fundamental logical principles of *modus tollens* and hypothetical syllogism do not seem applicable to such hypotheticals

(i) If Mary does not know that Brown is dead, then Brown is dead. But Brown is not dead. So, Mary knows that Brown is dead.
(ii) If Mary does not know that Brown is dead, then Brown is dead. But if Brown is dead, then Mary knows that Brown is dead.

So if Mary does not know that Brown is dead, then Mary knows that Brown is dead.

(iii) If Brown is dead, then Mary does not know that Brown is dead. But Mary knows that Brown is dead. So, Brown is not dead.

(iv) If Brown is dead, then Mary does not know that Brown is dead. But if Mary knows that Brown is dead, then Brown is dead. So, if Mary knows that Brown is dead, Mary does not know that Brown is dead.

If the hypotheticals in question are not genuine logical compounds, then what are they? The type of hypothetical involved in (23) (as well as in (14)) seems to be one which expresses the kind of implication we defined above. In other words, (23) seems to express a claim that if one is to use the language non-misleadingly in saying 'Mary does not know that Brown is dead', then one must believe that Brown is dead. If this interpretation is correct, then since the hypothetical is not a genuine logical compound we would avoid the odd application of *modus tollens* and hypothetical syllogism illustrated above. Moreover, we would forestall the possible objection that if we were to make both of the statements

(23) If Mary does not know that Brown is dead, then Brown is dead
(14) If Mary knows that Brown is dead, then Brown is dead,

then, supposing that Brown is not dead, we would be committed to the absurd consequence that Mary both knows and fails to know that Brown is dead.

The type of hypothetical involved in (18) (as well as in (17)) seems to be one which expresses what can appropriately be called a *conditional assertion*—a statement in the making of which a speaker, rather than doing something which amounts to the assertion of a conditional (a genuine logical compound), does something which has the result that: (i) if the protasis is true, then the speaker is committed to the assertion of the apodosis, and (ii) if the protasis is false, then the speaker is not committed to the assertion of anything.[1] Conditional assertions are usually employed in ordinary discourse to avoid commitment to presuppositions. (Conditional assertions should be distinguished from assertions typically made with sentences such as 'If you are interested, Brown is dead', in the use of which the speaker unconditionally asserts the apodosis, while avoiding commitment to something which is not a presupposition, but which could otherwise be legitimately taken by the audience to be among the speaker's beliefs.[2]) If such hypotheticals as

---

[1] Cf., e.g., W. V. O. Quine, *Methods of Logic*, rev. ed. (New York: Holt, Rinehart and Winston, 1959), p. 12; Nuel D. Belnap, Jr. 'Conditional Assertion and Restricted Quantification', *Nous*, IV (1970), 1–12.

[2] For a discussion of the latter type of hypothetical, see J. L. Austin, 'If's and Can's', *Philosophical Papers* (Oxford: Oxford Univ. Press, 1961), 205–232, pp. 212, ff.

(18) are construed as conditional assertions, then, since they would not be taken to be genuine logical compounds, the odd applications of *modus tollens* and hypothetical syllogism illustrated above would be avoided.

This interpretation of such hypotheticals will, in addition, forestall the following possible objection. Suppose that we make both of the following perfectly compatible statements

(17) If Brown is dead, then Mary knows that Brown is dead
(24) If Brown is not dead, then Mary knows that Brown is not dead.

If these two statements were genuine logical compounds, then in view of the fact that Brown is either dead or not dead, we would be committed to the statement

(25) Mary knows that Brown is dead or Mary knows that Brown is not dead,

a statement which seems rather odd, and one which, on the presuppositional view being expounded here, would have conflicting presuppositions. (The first disjunct presupposes that Brown is dead, while the second disjunct presupposes that Brown is not dead.) If the hypothetical statements (17) and (24) are conditional assertions, however, then in view of the fact that Brown is either dead or not dead, it follows that we are either committed to asserting that Mary knows that Brown is dead, or committed to asserting that Mary knows that Brown is not dead, a situation which lacks the objectionable quality of the situation of being committed to asserting (25). This interpretation would also account for the fact that the joint assertion of (17) and (24) is simply an alternative way of asserting that Mary knows *whether* Brown is dead, and would permit us to relate statements of the form '*a* knows whether *p*' to those of the form '*a* knows that *p*' while avoiding such odd statements as (25).[1]

It appears, then, that there is substantial evidence that apparent logical compounds which fail to share the presuppositions of their components are not genuine logical compounds. Thus, the "syntactic" conception of presupposition provides us with a promising means of clarifying and explaining the phenomena we have been dealing with. In light of this new interpretation of presuppositional phenomena, we have a good case for the general thesis that statements containing the phrase '. . . knows that *p*' involve the presupposition that *p* is true.

9. We must now apply the general thesis to the case of Smith's statement

(1) I do not know that Brown is dead.

---

[1] For a related discussion of the kinds of hypotheticals mentioned in this section, see my 'Hypotheticals: Conditionals and Theticals' *The Philosophical Quarterly*, XXIII (1973), 335–345, pp. 343, ff.

None of our other examples have involved first person usage. Nevertheless, first person usages in tenses other than the present seem not to constitute exceptions. For example, if someone made the statement

(26) I did not know that Brown was dead

he would imply that Brown was dead. Similarly, if someone, thinking that someday he would forget that Brown is dead, made the statement

(27) I will not know that Brown is dead

he would likewise imply that Brown is dead.

What happens when a statement of the form '*a* does not know that *p*' is made employing the first person present tense? If the statement really does *presuppose* the truth of *p*, then a speaker making the statement would imply that *p*. Thus Smith would have implied that Brown is dead. Such an implication would make Smith's statement have conflicting implications. For in ordinary conversational situations a speaker who implies that *p* normally implies that he knows that *p*. In other words, normally if a speaker implies that *p* and yet does not believe that he knows that *p*, then he is using the language in a misleading way. A speaker who wishes to express a belief that *p* when he does not believe that he knows that *p* is expected to provide some indication of this fact. Now it seems that if Smith had believed that Brown is dead while believing that he did not know that Brown is dead, we would expect him to have said, e.g.,

(3) I do not *know* that Brown is dead.

Since Smith (by hypothesis) employed instead the stress-free form of statement, he had not made use of the standard linguistic devices which indicate that one believes but does not know. Hence, we can say that he has indicated to his audience that he not only believes, but knows, that Brown is dead, an implication which conflicts with the assertive content of his statement. Thus, if Smith, in making his statement, really did imply that Brown is dead, since he failed to employ the standard linguistic devices discussed above, he also implied that he knew that Brown is dead. Nevertheless, the question remains: Did Smith, in making his statement, really imply that Brown is dead?

10. I now wish to claim that there is sufficient evidence that Smith did imply that Brown is dead, and hence made a self-stultifying, rather than a self-certifying, statement. As evidence I cite the previously noted features of logical compounds of statements of the form '*a* knows that *p*'. It would seem *ad hoc* to claim that only negative logical compounds employing the first person present tense fail to involve the implication

that *p*. And it would seem equally *ad hoc* to allege that negative compounds of first person present tense statements are not genuine logical compounds. As additional evidence I cite the apparent fact that we seldom, if ever, make or hear stress-free statements of the form 'I do not know that *p*' in ordinary conversational situations. We would not be surprised to hear Smith making the statement

(2) I do not know whether Brown is dead,

in which case he clearly would not be implying that Brown is dead. Nor would we be surprised to hear Smith making stress-containing statements (with the emphasis being supplied either vocally or by the context) such as the following

(3) I do not *know* that Brown is dead,

indicating that while he believes that Brown is dead, he does not wish to claim knowledge;

(4) *I* do not know that Brown is dead,

indicating that even though he does not know the facts of the case, there may be others who do;

(5) I do *not* know that Brown is dead,

indicating that he lacks knowledge someone has claimed or implied that he possesses. But it appears rather odd that Smith should make the stress-free statement

(1) I do not know that Brown is dead.

If this statement does in fact involve the implication that Brown is dead, then, as has been shown, it would involve conflicting implications, and hence its oddity would be accounted for.

Thus it appears that there are good reasons for maintaining that Smith's statement involves the implication that Brown is dead, and hence is self-stultifying rather than self-certifying. This is, as we have seen, a consequence of the fact that knowledge *presupposes* truth.[1]

---

[1] I would like to express my appreciation to my colleagues and students at SIU-E for helpful discussions on this topic.

Southern Illinois University at Edwardsville

WHAT MAGELLAN'S VOYAGE DIDN'T PROVE OR WHY THE EARTH IS FLAT

By JERRY S. CLEGG

WHEN seeking to indict some belief as archaic, superstitious or ridiculous, the knowledgeable polemicist is likely to compare the position he opposes to the belief that the earth is flat. He will do this on the presumption that scarcely anything could serve as a better example of a falsehood which only the ignorant or the unthinking have accepted. Magellan's voyage proved that the earth is round. Therefore, so we are all prone to conclude, the earth isn't flat. Those who say it is are to be corrected if teachable, dismissed if not. In truth, however, the earth is flat, more or less. The conviction that it isn't affords a fine example of how easily and powerfully proof can be misconceived.

A typical "demonstration" of the folly of believing in the earth's flatness will begin by pointing out the difference between something round, such as a tennis ball, and something flat, such as a table top. The question will then be asked which of the two things the earth is like. Since we know that the earth is a sphere and has no edges, we will say it is like the ball. After saying this we are apt to conclude that it, like the ball, is dissimilar to the table top, and so not flat. The obvious contrast between the ball and the table top appears to give us no other option. We are likely to feel certain that because the earth is round it just can't be flat.

Persuasive as it is, a "demonstration" of this kind is faulty. Implicit in it is the assumption that flat and round are always contrasting features. Normally they are contrasting, and this explains our readiness to say they always are. Still, some objects are both round and flat. The earth is a case in point. To see that it is, one must bear in mind how in our practical affairs we use the word 'flat'. A horizontal surface we are prepared to call flat is one we recognize to be neither a dome nor a bowl. It is neither concave nor convex, but a level expanse. Our grounds for saying this have nothing to do with a measurement of zero curvature. Rather they pertain to the behaviour of liquids and solids placed on its surface. We call a table top flat, for instance, because when a spirit-level is placed on it a reading of zero inclination appears. It does not shed water drops and marbles like an umbrella, nor does it collect them like a bucket. Any surface which does collect water at its centre is a bowl; any surface which does shed water is a dome. By the conventions of language, then, a flat surface is a level expanse on which solids and liquids tend to remain stable.

On these grounds there are plenty of regions on the earth's surface which are flat. Its regs, mesas, prairies, soda pans, lakes and oceans certainly come very close to being what any plain spoken man interested in having himself understood would call flat. One walks neither uphill nor downhill when treking across the Utah salt flats. They are, then, appropriately named. They match, too, the general curvature of the earth. They, and all similar areas, illustrate the point that being flat and being curved are compatible. Indeed, a little thought must show that a flat surface is necessarily a curved surface. What keeps a marble from rolling off a table is the uniformity of gravitational attraction at all points along the surface of the table-top. Since the force of gravity varies with the distance from the centre of gravitational attraction, all points on a horizontal, flat surface are equally distant from some one point. For us that point is approximately the centre of the earth. Because a surface can have all its points equally distant from some one point outside itself only if it is curved, flat surfaces must be curved. All their points must be equally distant from a point close to the earth's centre. Although variable, the earth's surface in general meets this condition. Its curvature acts as what might be called a natural paradigm of flatness. Surfaces whose curvature approximates that of a level expanse of the earth just are those we call flat. Any arc greater than the earth's curvature—such as that of a tennis ball—will result in a dome from which liquids and solids will flow, slide or roll. Any less an arc—such as that of a bucket—will result in a basin or bowl into which water and debris will collect. Flatness, so to speak, is a certain degree of curvature approximating that of the earth itself. If the right hand side of a perfectly flat table were to be extended, it would eventually curve around the earth to join itself on the left. We do not make our tables this large, of course, and so they all have edges from which items will fall if pushed far enough. This explains why, when mentally comparing a table-top to the earth, we tend to think of both as having edges. That is our true mistake, not the thought that the earth's surface is flat.

In case this observation seems dubious, a thought experiment will confirm it. Imagine the earth with one of its polar regions scraped off to 50 degrees of latitude so that its surface there parallels a ray of light. Many people would be tempted to say the earth would then be flat in that region, just as many now say the poles are, in fact, comparatively flat because their curvature is less than what we find at the equator. The temptation to speak this way stems from our entertaining a faulty picture of a flat expanse as a surface of zero curvature. It is a temptation which experience would very quickly dissipate. If left standing in the centre of a scraped off polar region, we would note immediately that we were in a huge depression. Any walk we took would lead us further away from the earth's centre against the pull of gravity. Any walk would, then, be a climb, and no one who had to hike to the rim of the depression would consider the territory he crossed flat. At 50 degrees of latitude the slope would approach some 20 degrees of inclination.

Clearly, a flat surface is one whose curvature approximates the general curvature of the earth. Those who think otherwise are guilty of the same mistake our medieval ancestors made. Some of the men sailing with Columbus feared they would sail off the earth, thinking it to be flat and with edges. We now discount those fears on the ground that the earth is round and without edges. The fear and its dismissal, however, rest on the shared error that what is flat cannot be round. The terror of the medieval sailor and the equanimity of mind of the modern seaman have in common a faulty picture of what it is for a surface to be flat. An increase in geographical knowledge has dispelled a groundless fear, but it has not corrected a faulty conception. Indeed, in one respect our knowledge has added to a store of confusion. From an essentially correct belief in the flatness of the earth the pre-Columbian sailor drew a wrong inference on the danger of falling into space, but at least there was no discrepancy between the truth, his usage and his conviction on what is flat. What he called flat is what he rightly believed to be flat. The modern sailor who erroneously denies that the world is flat faces, however, the embarrassment that his usage and his convictions are inconsistent. He is almost certain to call anything whose curvature matches that of the earth flat, even though on reflection he will deny that what he is prone to say is true. He is not in a position to "correct" his normal habits of speech either, for if he tried he would find that there is nothing he could call flat and still be understood. His offering a surface of zero curvature as an example of what he means would only puzzle others, for what is the point of calling a basin flat?

From this example of how easily proof can be misconceived a lesson can be learned. It begins with the observation that a more or less spontaneous picture, or imaginative conception, can stand in the way of our seeing how we use a word. The picture of an uncurved surface gives us a wrong idea of our use of the word 'flat'. The lesson continues with the warning that in such a case we should not follow the advice of pragmatists and positivists on how to clarify our ideas by asking if reality corresponds to our picture. Imagining what would verify or refute the thesis that the earth is flat, for example, is apt to mire us deeper in a muddle, leaving our faulty picture intact but our convictions now fully at odds with our usage. For all of us in our unreflecting moments nothing could serve as a better example of a flat surface than a calm lake or salt deposit. Yet, when invited to think about the matter, we are prone to conclude that Magellan's voyage proved such expanses to be less than flat. The imagined "proof" is deceptive; it is the unreflecting usage which is right. Things, after all, are what we say they are. Clarity of mind requires that we honour our usage and resist the "proof" which only stengthens the force of a faulty picture.

Mills College, Oakland, California

THE EXAMINER EXAMINED

By B. H. SLATER

TO untie knots, whether in string or in the mind, one must first loosen the material.

'Pupils will have an exam one day next week. Pupils will not know in advance which day it will be.'

Pupils hearing these words may argue 'Well, it can't be on the last day or else we would know on the previous night. But then it can't be on the penultimate day either, because by the night prior to that it could only be on the next day or the last. The last is ruled out, so it would have to be the next to last; and that means that once again we would know in advance which day it would be. But we could go on like this right the way back through the week. So when can it be? There is no time it could be, if the notice is correct. The teacher is contradicting himself. So there may or may not be an exam any time during the week.'

The teacher, when he in fact holds the exam on the Tuesday, may argue 'Well, here it is, the exam, on one of the days of the week. So my first remark is correct. And none of you was confidently and with reason expecting it today. So my second remark was correct. I'm not contradicting myself.'

What is going on here? Is the pupils' argument valid? Does the teacher vindicate himself? Consider the following parallel.

'Pupils will have an exam on Tuesday. Pupils will not know in advance which day the exam will be on.'

Pupils hearing these words may argue 'If we are told we are going to have an exam next Tuesday, the date of the exam can be foretold, it is expected, it's no surprise. What the teacher says is self-contradictory. So there may or may not be an exam next Tuesday. We have no way of telling for sure.'

The teacher, when he holds the exam on Tuesday, as he promised, can say 'Well, here it is at noon on Tuesday, the exam. So my first remark was correct. And none of you was confidently expecting it; so my second remark was correct. Both things I said are now shown to be true; so what I said was not self-contradictory.'

Clearly there is no conflict between the exam's being on Tuesday and the exam's being unexpected. The conflict lies between the pupils' *knowing* when the exam will be and the exam's being unexpected. 'Pupils will not be expecting the exam' is contrary to 'The pupils know the exam will be on Tuesday', not to 'The exam will be on Tuesday'. So what the teacher *says* is quite consistent.

The pupils' reasoning is therefore invalid. Trusting the teacher, they argue 'Teacher says so, so we know it is so. We know there will be an exam on Tuesday. Yet the exam will be unexpected by us. What a mix up! Will it be on Tuesday, or not?' The pupils learn to expect an exam, not from some statement given, but from the fact that they have been given some statement.

So the behaviour of the teacher is equally reprehensible, although in a different way. He says 'There will be an exam on Tuesday', but he also denies that the pupils can know this by saying in addition 'The exam will be unexpected'. While not making contradictory remarks, by *making* the first remark in his position of authority he influences the truth of the second. It is not so much that the truth of the first opposes the truth of the second, but that his asserting the first makes the second untrue.

The essential point here goes also for the case where the teacher is less specific about the day of the exam. What is at fault is the thought that we can infer that we know the exam will be one day during the week from 'The exam will be one day during the week'. There can be no conflict between the latter and any statement about expectations; but someone's saying the latter can influence expectations.

Consider another parallel. Suppose someone says to us 'My name is Tom; but you don't know what my name is'. There is no conflict between his name being Tom, and our not knowing his name. So his statements are not contrary. It is his performance in telling us his name that makes impossible our remaining in ignorance of it. 'We don't know what your name is' is an inference from what has been said. 'We have been told your name is Tom' is a description of what has been said. It is these two which are at odds: one can't both not know, and at the same time have learnt, what his name is. But 'We learn, from what you say, that your name is Tom' is not *deduced* from what is said. It is something we say on the basis of observation, not inference. We learn by listening, not by arguing.

The way is now clear for a more exact examination of 'The Examiner'.

University of Kent at Canterbury

WHITELEY ON RULES OF LANGUAGE

By C. R. CARR

C. H. WHITELEY (ANALYSIS 34.2, pp. 33–38) proposes two arguments for the impossibility of rigidly circumscribed semantic rules, rules which 'tell us what to say if we wish to convey a given sense', 'which determine for each sentence whether it is significant or not and whether it has one or many senses . . .'. These quotations suggest that his target is the sound-meaning pairings of linguistic semantics. I would like to argue that this isn't Whiteley's target, and that his arguments fail because he does not distinguish two different tasks. As a result he is attacking one quarry with arguments designed for another. The failure to distinguish these two tasks is underscored by an equivocation in his arguments.

The first argument notes that actual linguistic performance contains slips of the tongue, mispronunciations and related errors. Whiteley correctly claims that we still know in many of these cases what the utterance means even though this is different from what the rules for the sentence say the sentence means. A husband preoccupied with bills intends to ask his wife to turn down the radio, but says 'Turn down the debt'. She understands that he wanted her to turn down the radio. In such cases one can't help but understand what is being said. As Whiteley notes, 'Whether I make sense of an utterance is not a matter of choice'.

'Utterance' in this argument is ambiguous. In asking what an utterance means we might be asking what the speaker means by uttering the words or what the words uttered themselves mean. What semantics seeks to do, and what the quote about determining senses suggests that Whiteley is criticizing, is relate utterances in the second of these senses to meaning in a systematic way. Nothing he has said has shown that semantic rules which describe such a sound-meaning pairing can't exist. Indeed, in cases such as those he has described, what we say to the person is something like 'You didn't mean that', i.e., what those words mean. How, if sentences aren't related to meaning in a systematic way, are we able to say that a particular sentence could not have been what was meant? His argument presupposes the existence of the very semantic rules he denies exist.

Whiteley's second argument claims that, just as we can't set limits to the variety of possible experience, so likewise we can't set limits to the variety of intelligible utterance. We need an 'open-texture' for meaning. Sometimes only by metaphor can we adequately say what we mean. Strict meaning rules would, he believes, preclude the use of metaphor.

A typical case of metaphor involves the speaker wishing to convey something more than his words convey or something slightly different from what his words convey. There are two reasons for rejecting the suggestion that these uses of language are incompatible with strict meaning rules. First, not only are semantic rules compatible with metaphor, but without such rules the distinction between metaphor and more literal uses of language would collapse. A metaphor is an extension of the meaning of a sentence, but this requires that the sentence be such that it is definitely related to some meaning or set of meanings. If not, how does a metaphorical use of language differ from a non-metaphorical one? The answer can't be that literal utterances are those that are more often or more generally associated with a certain meaning. Since most of the sentences in the language which are produced are uttered only once it is hard to see what makes them the ones standardly or generally associated with a certain sense, if not the fact that they are related to that sense by certain rules. And a metaphor which is widely used does not necessarily cease thereby to be a metaphor, which it should do if literal meaning is just being generally associated with a certain sense.

The second reason for rejecting the argument from metaphor is the same as the reason for rejecting his argument from performance mistakes. Unless sentences are in some strict systematic way related to a meaning, it makes no sense to say of a given utterance that it is an extension of the meaning of that sentence. Semantic rules give the best account of that relation.

Whiteley has done nothing to show there cannot be strict semantic rules in the sense in which semanticists talk of such rules. His very arguments in fact depend on the existence of such rules. What he is really arguing against, although his introductory comments obscure this, is not semantic theory as it is usually understood. He is denying the possibility of a more general theory for finding out in a given case what a person's intentions are when his actions aren't sufficient to determine this. That is, he is saying there can be no rules by which the wife can determine that her husband wanted her to turn down the radio from the words 'Turn down the debt'. Many of those he purports to be criticizing would agree. But clearly his arguments do not show that such a theory is impossible, but only that a theory which wasn't designed to do the job can't do it. It is like arguing that a theory of physics is impossible because the conceptual equipment of biology can't account for the phenomenon of gravity. That rules designed to show 'for each sentence whether it is significant or not . . .' cannot tell us what a person meant when he accidentally used the wrong word does not show that there are no rules which can tell whether a sentence is meaningful and what it means. Nor does it show that there cannot be a theory of what a person meant when this is underdetermined by what he said. It will be a different theory from the former, and will have to include items like background information, shared beliefs and context. To show that such a theory is impossible will involve showing that not even a theory that makes reference to these items can explain how we determine what a speaker meant.

We do sometimes tell when what was said isn't what was meant. Why, a priori, believe that a theory of what we do is impossible?

University of Arizona

'IS MORALITY A SYSTEM OF HYPOTHETICAL IMPERATIVES?' A REPLY TO MR. HOLMES

By PHILIPPA FOOT

IN ANALYSIS 34.3 Robert Holmes criticizes 'Morality as a System of Hypothetical Imperatives'[1] where I argue that moral judgements are not categorical imperatives. Am I, he wonders, denying the intelligibility of the categorical/hypothetical distinction, so making nonsense of my own title? The answer to this is that although I find unintelligible some of the things people say in trying to separate moral injunctions from hypothetical imperatives, as for instance that they have a special necessity or binding force, there is one version of the doctrine which seems to me to rise to the level of falsity. I mean the idea that moral judgements, unlike hypothetical imperatives, have an automatic reason-giving force. As Frankena put it, quoting from G. C. Field, '. . . it is "one of the most deeply recognized characteristics of the moral fact" that it is in itself and necessarily "a reason for acting".'[2] This is what I am denying when I deny that moral judgements are categorical imperatives: in my view moral considerations give reasons for acting only when relevantly related to the agent's interests or to his (altruistic and non-altruistic) desires.[3] Mr. Holmes is prepared to accept this, the rather obvious, interpretation of my article, but thinks that it must be a trivial point that I am making, hardly one from which so momentous a conclusion could be derived. For of course I cannot be denying that in *one* sense of the word 'reason' moral considerations necessarily give reason for acting. I seem to have overlooked the distinction between 'justifying reasons' and 'motivating reasons', and moreover to have forgotten that reasons give reasons from different points of view. What a moral consideration necessarily gives is, says Mr. Holmes, a justifying reason from the moral point of view: no one judging the omission of an act to be morally bad can deny that he has 'a morally sufficient reason for doing that thing' (p. 99).[1]

---

[1] *Philosophical Review*, 1972, pp. 306–316.

[2] W. K. Frankena, 'Obligation and Motivation', in A. I. Melden (ed.), *Essays in Moral Philosophy*: Univ. of Washington Press, Seattle.

[3] Whether a man necessarily has reason to do what it is in his interest to do is a matter of controversy. Nothing that I say in this rejoinder depends on this point.

Let us consider these points separately, and first the distinction—spoken of as if it were well established—between 'justifying reasons' and 'motivating reasons'. So far as I know this piece of terminology was given its present currency by Frankena, in the article already cited, and I find the same confusions in Frankena and Holmes. There is confusion, first, between giving reason for a judgement (e.g. a moral judgement) and giving reason for an action (the one mentioned in the judgement), and secondly between justifying the judgement and justifying the action. By using 'justifying reason' indiscriminately for each of these things it is made to look as if they were inseparable. But I had said that the reasons that establish the moral judgement may not establish a reason for acting. Mr. Holmes is simply begging the question when he uses the same term 'justifying reason' for 'a reason to support the judgement *that* one should do x' (p. 98), i.e. 'a justification of a moral judgement' (ibid.), and also for 'a justification for following that course' (ibid.) and a 'reason for doing that thing' (p. 99).

That the transition from 'reason for the judgement' to 'reason for the action' is illicit, and in the context question-begging, is obvious. The manoeuvre which involves *action justification* needs a little more sorting out. Here one moves from the justification of the judgement to the justification of the action, and from there to reasons for doing the action. Once again the transition is illicit, for either justifying the judgement does not count as justifying the action, or else justifying the action is not necessarily showing reason for doing it. That it is not immediately obvious where the break comes is due, partly, to some complications connected with the meaning of 'justifying an action'. They are not really important here, but I must try to get them out of the way.

In the first place only some judgements of morality, law or etiquette (which are all examples used by Mr. Holmes) could possibly be taken as providing justification for actions. Strictly speaking justification belongs where there has been some objection which the judgement overrules. And while no unfavourable judgement could give an act justification, it is not clear that a justification implies that law, morality, etiquette or some other such system is positively in its favour. It seems enough for justification if the action is shown to be innocent according to that system—always supposing that some objection appeared to lie. So, if a policeman finds me climbing through my neighbour's window, or destroying his property, the neighbour's prior permission justifies my action in the eyes of the law.

---

[1] I do not follow Mr. Holmes in taking 'ought' statements as my examples of moral judgements because I see some unresolved philosophical issues here. It is possible that 'ought' does imply reasons for acting, and if so I would challenge the propriety of making these particular moral statements unless such reasons have previously been established.

These considerations suggest that those who talk about justifying reasons, and want to link them with the justification of actions, are operating with a somewhat artificial version of this concept. But this does not matter, except in so far as it makes it more difficult to see things clearly. To bypass such difficulties I shall use examples such that there, if anywhere, the justification of a judgement would also be the justification of an action, and the justification of the action would provide a reason for doing it. I shall suppose that the judgement is strongly favourable to the action; it might even say that it would be morally wrong, illegal, or not *comme il faut* to leave the action undone, so that it would actually be required by morality, law or etiquette. And to make it natural to speak of justifying the action I shall suppose that it was alleged to be objectionable from one of these points of view.

I said that there must be a break either between justifying the judgement and justifying the action or between justifying the action and showing that there is reason to do it. Where does it come? Perhaps it does not matter very much whether we bracket the justification of the action with the justification of the judgement or with the giving of reasons for doing it. But the following example suggests that it is more natural to do the latter. We start with a man who denies, as many people do, that he has reason to do something simply because it is required by law. We suppose that he is told that he is doing something illegal, but that in fact it is not illegal: he would be able, if he chose, to show that the action was legally unobjectionable, or even, in the circumstances, required by law. The relevant question from our point of view is as to whether, if he were really indifferent to considerations of legality, he would describe himself as *justifying* his action? It seems to me that this would be strange language to hear from a hardened criminal, unless he were putting on an act before the police. To talk about justifying an action is, to imply reasons on one side or the other, and this, by our hypothesis, is being denied.

Let me turn now to the other concept which Mr. Holmes introduces in asserting that moral judgements necessarily give reasons for acting. He had said that they give justifying not motivating reasons and he adds that they give reasons *from a moral point of view*. This he supposes that I cannot deny—that moral considerations give reasons from a moral point of view—and thinks I must be making the trivial point that they need not give reasons from some *other* point of view. Before I can say whether or not I deny that moral considerations give reasons at least from the moral point of view, I must ask about the meaning of the expression 'reason from the P point of view'. I want to ask 'Is a reason from a point of view a reason?' meaning 'Is the assertion that I have a reason from point of view P to do x inconsistent with the claim that I have no reason for doing it?' If the answer is that it is inconsistent it will be proper for some to deny that they have any reason from the point of view of etiquette to do 'what's done', and others will say that they do not have reason from the point of view of law to do what the law requires. The same will go for morality unless my challenge—to show how this is different—is met. And this is a challenge that Mr. Holmes does not take up. Indeed he is just as insistent that there are reasons from the point of view of etiquette to follow the rules of etiquette as he is that there are reasons for acting morally from the moral point of view.

What I am suggesting is, of course, that either one may accept 'reason from a point of view' but deny 'reason', or else one may deny 'reason from a point of view'. The interesting question is whether in accepting 'reason from a point of view' and denying 'reason' one is, as Mr. Holmes suggests, merely denying a reason from some *other* point of view. The implication of such a suggestion would be that all considerations can equally be denied the status of reasons. And it is this implication which seems to me absurd. For suppose we do say that all reasons are reasons from a point of view and apply this not only to 'moral reasons', 'reasons of law' and 'reasons of etiquette' but also to reasons based on the agent's desires. We invent a name for this last 'point of view' and imply that there is nothing odd about denying that a man has reason to do the things that will get him what he wants. But what in fact would we say of someone who wanted to get out of a miserable situation, thought that action A would get him out of it, but denied that he had any reason, even other things being equal, to do A? Or suppose him to want something for someone else, and yet to deny that he had any reason at all for doing what would get it for him. That we should call such men irrational is just one pointer to the closeness of the connexion between reason and desire. Another, from the same network of concepts connected with the explanation of human conduct, is to be found in the conceptual connexions between motives and reasons. Whether or not reasons imply motives it seems that motives *by and large* imply reasons. And as one has a motive for doing an action that will further desired ends it follows that *by and large* one has reasons for doing these actions. One of these connexions must be broken by those who think it possible to sever reason and desire.

I conclude that in denying that considerations of law, morality or etiquette necessarily, and in themselves, give reasons for acting I am not making a merely trivial point. This does not matter very much, except in so far as concerns the survival of the worn out doctrine of the categorical imperative. I should like to see that disappear from the scene, and I wish it would take with it the 'justifying reason', a shady character of no fixed address.

Somerville College, Oxford, and The University of California, Los Angeles

ON ANSCOMBE'S EXPOSITION OF HUME

By A. C. GENOVA

IN her recent article (' "Whatever Has a Beginning of Existence Must Have a Cause": Hume's Argument Exposed', ANALYSIS 34.5, pp. 141–151), G. E. M. Anscombe misconstrues what Hume means by an abstract or general idea. In respect to Hume's proof that 'Whatever begins to exist must have a cause of existence' is not an intuitively or demonstrably certain principle (*A Treatise of Human Nature*, Book I, Part II, Section III), she contends that there is (perhaps) a sound formulation for this 'argument from imagination' but that it does not yield the conclusion Hume needs. I want to show that it *does* yield the appropriate conclusion if one correctly understands Hume's doctrine of abstract ideas (Book I, Part I, Section VII). Although Anscombe's paper also deals with several related issues (which I think also turn on the interpretation of abstract ideas), I shall restrict my discussion to what she rightly claims is the 'main business', viz. Hume's specific argument purporting to show that it is not absurd or contradictory to suppose the real possibility (and actuality) of an uncaused beginning of existence.

First, I think it would be helpful to be clear about Hume's general strategy concerning the controversial argument. It seems, *prima facie*, that what Hume is generally trying to establish is quite straightforward: If it is logically and psychologically possible to conceive, imagine or suppose that something could come into existence without a cause, then the purportedly necessary 'general maxim of philosophy, *that whatever begins to exist, must have a cause of existence*' (Selby-Bigge's edition, p. 78) is not a logically necessary truth and the supposition of an actual uncaused beginning of existence is a logical possibility (we do not thereby incur any absurdity or contradiction); and if the general maxim is indeed contingent, then any postulated 'necessary connection' between an object (or event) and some cause must be relegated to the realm of 'matters of fact' (not 'relations of ideas'), in which case, the relation of necessary connection must be locatable as an empirical datum which, according to Hume, is notoriously (but not surprisingly) absent when we look for it. Now the crucial step in this general strategy is the move Hume makes from imaginability to possibility, viz. the piece of argument which Anscombe criticizes. Let us briefly review her central criticism and then examine it in the light of Hume's doctrine of abstract ideas.

I

Again then, Hume's argument challenging the necessity of 'whatever begins to exist must have a cause of existence':

> But here is an argument, which proves at once, that the foregoing

> proposition is neither intuitively nor demonstrably certain . . . that as all distinct ideas are separable from each other, and as the ideas of cause and effect are evidently distinct, 'twill be easy for us to conceive any object to be non-existent this moment, and existent the next, without conjoining to it the distinct idea of a cause or productive principle. The separation, therefore, of the idea of a cause from that of a beginning of existence is plainly possible for the imagination, and consequently the actual separation of these objects is so far possible, that it implies no contradiction or absurdity. (Hume, pp. 79–80.)

Now Anscombe's difficulty arises in the context of deciding how to construe the denotation of the term 'objects' as it occurs in the last sentence of the quoted passage. Its denotation must obviously conform to the respective denotations of the antecedent terms 'a cause' and 'a beginning of existence' in the same sentence, but now we have an alternative in respect to how we are to interpret *these* indefinite singular terms and the alternative supposedly results in a dilemma. The alternative is to interpret the indefinite singulars as designating abstract ideas or particular ideas. Anscombe argues that it would be both reasonable and conformable to Hume's doctrine of abstract ideas to supply particular cases (or ideas) as the denotations of these terms. Accordingly, for 'a beginning of existence' we should supply a specific case (like the idea of a particular rabbit), '. . . for neither in reality nor according to Hume can there be a bare image of a beginning of existence which is not the beginning of existence of anything in particular' (Anscombe, p. 149). With that settled, she then poses the alternative (mentioned above) for the interpretation of the other term, 'a cause', i.e. we can similarly provide a specific case (like the idea of a particular parent rabbit) or forsake Hume's doctrine by letting the indefinite 'a cause' stand as a sufficient description of our idea.

The resulting dilemma is as follows: On the one hand, if we take the former course and conform to Hume's doctrine of abstract ideas, then the argument from imagination is sound because we can indeed imagine that any particular beginning of existence can occur without any particular pre-designated cause, which in turn implies that the supposition that this actually can occur is not absurd or contradictory. But *this* just establishes that any particular effect (or thing) can occur without any particular cause which you have imagined it without; it does not establish what Hume needs, viz. the general proposition that it is possible to imagine a beginning of existence without any cause at all, much less that such a possibility can be supposed to occur in reality without contradiction or absurdity (Anscombe, pp. 149–150). But on the other hand, if we take the latter course and 'forget' about Hume's ban on abstract ideas, then Hume's argument would be 'very unconvincing'. Why?

> The trouble about it is that it is very unconvincing. For if I say I can imagine a rabbit coming into being without a parent rabbit, well and

> good: I imagine a rabbit coming into being, and our observing that there is no parent rabbit about. But what am I to imagine if I imagine a rabbit coming into being without a cause? Well, I just imagine a rabbit coming into being. That this *is* the imagination of a rabbit coming into being without a cause is nothing but, as it were, the *title* of the picture. Indeed I can form an image and give my picture that title. But from my being able to do *that*, nothing whatever follows about what is possible to suppose 'without contradiction or absurdity' as holding in reality. (Anscombe, p. 150.)

II

What Anscombe misses, I think, is that Hume has already slipped between the horns of her dilemma. She seems to assume that, for Hume, there are only two exclusive possibilities: (1) that we can only think, imagine or conceive in terms of particular ideas or concrete images, or (2) that we can also perform these mental operations in terms of abstract ideas or 'bare images' lacking all reference to or association with particularity. She also assumes that Hume would countenance (1) and eschew (2). In any case, neither the implementation of (1) nor that of (2), in Hume's argument, yields his desired conclusion. I think that Anscombe is entirely correct about *that*, and I think Hume would agree. Hume does not merely want to prove the possibility that some specified particular can come into existence independently of its relation to some other specified particular; and he most certainly would not espouse the view that the mere fact that we can form a single image of something thereby proves the possibility of that something's existing uncaused.

What Anscombe overlooks is that Hume's position on abstract ideas is not (1) simply, as opposed to (2), but (3) that we *can* think in terms of abstract ideas (and typically always do so in our discursive reasonings)—not in the *traditional* sense of 'abstract idea'—but in a special more complex sense which is both compatible with the fact that every mental image must be particularized and completely determinate *and* the fact that our abstract reasonings continue to fulfil all the requirements of generality which characterized the traditional interpretation. The point is that Hume actually holds a *middle position* between (1) and (2)—an interpretation of abstract ideas which construes them as *particular* in their individual nature but *general* in their representation: 'Abstract ideas are therefore in themselves individual, however they may become general in their representation. The image in the mind is only that of a particular object, tho' the application of it in our reasoning be the same, as if it were universal' (Hume, p. 20). To be sure, Hume does insist upon the particularity of mental images, but this is only *part* of Hume's thesis about abstract ideas, and not the most important or philosophically interesting part. But how can Hume hold a view which seems to conjoin two incompatible elements (particularity and universality) in a single idea?

This is precisely the problem which Hume confronts in Book I, Part I, Section VII. Hume says flatly ' 'Tis evident, that in forming most of our general ideas, if not all of them, we abstract from every particular degree of quantity and quality, and that an object ceases not to be of any particular species on account of every small alteration in its extension, duration and other properties. It may therefore be thought, that here is a plain dilemma . . . concerning the nature of these abstract ideas . . .' (Hume, p. 17). The 'dilemma' presented by his hypothetical *opponents* is precisely that 'The abstract idea of a man represents men of all sizes and all qualities; which 'tis concluded it cannot do, but either by representing at once all possible sizes and all possible qualities, or by representing no particular one at all' (Hume, p. 18). Hume clearly wants to support a notion of abstract ideas which *does* 'represent men of all sizes and all qualities' but requires neither the representation of all possible determinations at once nor the representation of none. The former requirement is immediately ruled out because of the finitude of the human mind; the latter requirement (which, incidentally, is what Hume thinks the traditional view presupposes) is ruled out on two counts: (a) by three arguments (Hume, pp. 18–20) designed to prove that it is impossible to conceive of any quality or quantity without forming a precise notion of its degrees, and (more relevant to my purpose) (b) by showing that although the mind is not infinite, nevertheless, 'we can at once form a notion of all possible degrees of quantity and quality, in such a manner at least, as, however imperfect, may serve all the purposes of reflexion and conversation' (Hume, p. 17). In other words, Hume is saying that his (and Berkeley's) reformulated interpretation of abstract ideas can do everything the traditional view could do, only better. An abstract idea can represent all possible determinations of the object and still be no more than a particular idea in the mind. But how?

Well, Hume's explanation need not be dwelled on. The important thing to grasp is that he holds such a position, tenuous though it may be. He says that when we find a 'resemblance' among particular objects, we can apply the same name to all of them regardless of their other differences. After we acquire a *custom* or habit of this kind, the recognition of the name calls up the idea of one of the particular objects and we imagine it in its specificity. But since the same word is also associated with and applicable to the other resembling ideas which differed in specificity from the particular idea presently before the mind, the word (not being able to revive each and every associated idea) nevertheless 'touches the soul' and can revive the custom or habit we originally acquired in surveying the multiplicity of resembling ideas. Thus all the resembling ideas are *potentially* present to the mind so that we thereby keep ourselves in readiness to survey *any* of them as may be required. 'The word raises up an individual idea, along with a certain custom; and that custom produces any other individual one, for which we may have occasion' and the 'extraordinary circumstance' is that 'after the mind has produc'd an individual idea, upon which we reason, the attendant custom, reviv'd by the abstract term, readily suggests any other individual, if by chance we form any reasoning that agrees not with it' (Hume, pp. 20–21). The upshot is that a particular idea *becomes* general by being annexed to a general *term* with which it has a customary conjunction and thereby achieves a reference to *any* other resembling particular idea *qua* resembling. Thus, in this context, any reasoning which holds for one holds for all. That is how our reasoning achieves generality for a nominalist like Hume.

When we apply this exposition of Hume's doctrine of abstract ideas to Anscombe's exposition of Hume's subsequent argument concerning the possibility of an uncaused beginning of existence, we see that *both* of the terms 'a beginning of existence' and 'a cause' stand for abstract ideas in Hume's sense. For he wants to prove the general conclusion that since all which is relevant to the abstract idea of a beginning of existence (that is, any object just insofar as the general term 'a beginning of existence' applies) is separable in the imagination from all which is relevant to the abstract idea of a cause (that is, any object just insofar as the general term 'a cause' applies) there is no contradiction or absurdity in supposing that *any* of the objects respectively associated with each of these abstract terms are also separable in reality. But *that* is just to say that an uncaused beginning of existence is a logical possibility for the real world.

So for Hume an abstract idea is not a 'bare image' of nothing in particular because we have seen that that interpretation is exactly what Hume is rejecting. Nor do we 'forsake' Hume's doctrine by construing the relevant terms as standing for abstract ideas because we have seen that Hume is defending an interpretation of abstract ideas. And finally, it is not true to say that when we imagine a particular rabbit existing without a cause, we are thereby just imagining a particular rabbit existing. Rather, we are comparing, in our imagination, the particular idea of a determinate rabbit with the general idea of a cause (the idea of a 'productive principle' for Hume); and we notice that the former idea can be imagined independently of all that pertains to the latter idea—just as the idea of a body (abstract or particular) is likewise separable from the abstract idea of a mind, but not from the abstract idea of extension. The trick is to grasp Hume's ingenious device of characterizing an abstract idea as a complex consisting of a particular idea which represents all possible determinations of resembling particulars *in virtue of* its association with a general word whose occurrence revives the habit of surveying all or any of the resembling particulars *qua* resembling in the relevant respects.

My defence of Hume is not of course meant to imply that Hume's position is unassailable. He makes everything turn on 'the manner' in which the faculty of imagination is ordered, and all ontological and epistemological differentiations are reducible, in the final analysis, to modes of imagination. One can also question his interpretation of necessity as exclusively logical necessity, thereby excluding any notion of an a priori synthetic necessity (which would not require logical inseparability) or an empirical necessity (which might be attributed to empirical laws even though they be logically contingent). But *these* are different issues. I have been arguing that Hume's argument for the possibility of an uncaused beginning of existence is sound if (as Anscombe presumably does) we analyse it in terms of Hume's empiricist framework and principles.

University of Kansas

WRIGHT ON FUNCTIONS

By PATRICK GRIM

LARRY Wright's 'Functions' (*Philosophical Review* 1973) is one of the most complete attempts to date to give a satisfactory account of sentences of the form 'The (or a) function of X is Z' and their explanatory force. But counter-examples are easy to come by and a clear philosophical outline of functions is still a long way off.

I

Wright summarizes his account, with the warning that we must understand the key terms as he has outlined them, by identifying the meaning of 'The function of X is Z' with that of the conjunction of the following

(a) X is there because it does Z
(b) Z is a consequence (or result) of X's being there.

He insists that the 'because' of the formulation is used in 'its ordinary, conversational, causal-explanatory sense'. He later notes that 'When we say the function of X is Z in these [conscious] cases, we are saying that at least some effort was made to get X (sweep hand, button on dashboard) where it is precisely because it does Z (whatever)'. But since he wants to present a unified account of natural and conscious functions this can only be taken as a simple illustration of how particular cases fit the formula-

26 DEC 1975

tion, and not as a criterion of the formulation itself. We are left with the bare "causal-explanatory" 'because' of the formulation itself.

Three readings are outlined for 'X is there': 'X is where it is', 'a C has an X' and 'X exists at all'. Wright further suggests that 'in some contexts we will allow that X does Z even though Z never occurs' and that ' "Z is a consequence of X" must be consistent with Z's not occurring'.

I will attempt to avoid the problems of 'do's that don't by structuring examples so that only full-blown 'do's and 'consequences' appear. I hope to show that Wright's account is insufficient for either natural or consciously designed functions, regardless of the reading chosen for 'X is there'.

II

A. The Chocomotive is running very roughly indeed owing to a large gap in the piston rings of the third cylinder. The excessive vibration of the engine works loose a pollution control valve, designed to save our lungs, but which makes the engine run a bit less smoothly. Located at the mouth of the carburetter, the valve is sucked through the intake manifold to cylinder number three. It there coincidentally lodges itself in the ring gap, making the engine run smoothly and efficiently once again. A roughly running engine would have shaken any piece of metal in cylinder number three out through the exhaust in a short while. But the pollution control valve smooths the engine's performance just enough for it to stay there indefinitely.

One day Browness and Amos notice that the pollution control valve has been sucked into the intake manifold and dismantle the engine for practice. Amos notices the pollution control valve nestling in the ring gap of the third cylinder and asks why it's there rather than embedded in the head or outside the exhaust as one might have expected. Browness can tell Amos that its making the engine run smoothly is a factor in its being there rather than elsewhere, and that it is there (at least in part)[1] because it makes the engine run smoothly. Both know that the engine's running smoothly is a consequence of its being there. But Browness would be wrong to tell Amos that the (or even a) function of this or any similar pollution control valve is to make the engine run smoothly.

Interestingly enough, the case of Browness and the Chocomotive can be construed so as to fit any of the readings offered for 'X is there'. The pollution control valve is where it is (rather than embedded in the head or outside the exhaust) because of what it does, this engine has one (rather than having lost it) because of what it does, and it may even exist at all (rather than being smashed to atoms by the nuclear reactor Browness uses for a silencer) because of what it does.

---

[1] A parenthetical qualification which Wright's examples, presenting neither necessary nor sufficient conditions, demand as well.

B. Wright's formulation proves similarly insufficient for natural functions:

Jumpin' Jeanne overwaters her one-leafed hydrangeas, with the consistent result that they fall into the puddles around their bases and drown. Drowning hydrangeas lose their leaves. One day Buck carelessly brushes against a particular plant's single leaf, wedging it against a neighbouring shrub so as to support the hydrangea and keep it from drowning.

Buck later asks Jeanne why this particular plant has its leaf, whereas the others have lost theirs. She can tell him that its keeping the plant from drowning is a factor in the plant's having any leaves at all, and that it has the leaf because it keeps the plant from drowning. She can further say that the plant's not drowning is a consequence of its having the leaf. But she'd be wrong to tell Buck that the (or even a) function of this or any other hydrangea leaf is to keep the plant from drowning.

The case of Jumpin' Jeanne, like that of the Chocomotive, can be construed in line with any of the suggested readings for 'X is there'. The hydrangea leaf is where it is (rather than in the puddle) because of what it does, the hydrangea has a leaf (rather than having lost it) because of what it does, and perhaps the leaf exists at all (rather than becoming mulch) because of what it does.

III

In many ways Wright's 'Functions' is a significant contribution to an intriguing and important study. But his more general claims, such as the definitional unity of conscious and natural functions and the 'etiological' nature of functional 'ascription-explanation', rest on his specific formulation and the insufficiency of its rivals. The formulation he presents, I think, has proved inadequate.

Boston University

NOTES

The Analysis Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Margaret A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, A. C. MacIntyre, Bernard Mayo, D. H. Mellor, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

Subscriptions. The subscription to Analysis is £2.50 ($7.50) for individuals in Great Britain, £3.00 ($9.00) for individuals abroad, and £3.50 ($10.50) for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford, or placed with any bookseller.

Contributions. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgment is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in Analysis remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

26 OCT 1975

PRINTED IN GREAT BRITAIN BY BURGESS AND SON (ABINGDON) LTD.

l. 35 No. 3 (New Series No. 165) January 197

ANALYSIS

Edited by
C. J. F. Williams

UNIVERSITY CALCUTTA

CONTENTS

Moral notions and moral misconceptions	Keith Graham
Indirect reference	Thomas Baldwin
Conditionals, impossibilities and material implications	Peter Downing
Further in the modest defence	Joel Rudinow
D. Z. Phillips and 'the inadequacy of language'	David A. Conway
Substitutionalism and substitutional quantification	R. D. Gallie
Gods and viruses	Norvin Richards
Utterer's meaning and implications about belief	Michael Clark
An observation on English truth-functions	Keith Halbasch
Analogy, induction and other minds	Theodore W. Budlong

BASIL BLACKWELL . ALFRED STREET . OXFORD

Ninety pence

ANALYSIS 35. 3 JANUARY 1975

MORAL NOTIONS AND MORAL MISCONCEPTIONS

By KEITH GRAHAM

IN *Moral Notions*[1] Julius Kovesi marshals, against a view of ethics to which I am unsympathetic, a set of lively, penetrating and original arguments to which I am equally unsympathetic. My ultimate aim in criticizing his criticisms is to draw attention to a kind of freedom, rather different from the "moral autonomy" talked about *ad nauseam* in recent years, which I believe exists in moral matters and which is obscured both by Kovesi and by his opponents.

I

Kovesi is unimpressed by the piece of conventional wisdom which says we cannot move from a set of descriptive premisses, to the effect that a thing or action has certain natural properties, to the evaluative conclusion that that thing or action is good. He is unimpressed not because he thinks the gap is easily bridged, but because he thinks there is as much of a problem in the majority of cases where we wish to move from such premisses to a *descriptive* conclusion (p. 9). In the terminology which he adopts, there is no entailment relation between the material and formal elements of a notion. Whether the thesis Kovesi is putting forward here is true, and if true interesting, depends to some extent on the clarity of the technical terms which he introduces. I shall suggest that these terms are unclear in a way which cannot be remedied: they embody a fundamental confusion.

We are told (p. 4) that the *material element* of a notion—the notion of, say, a table or an inadvertent act—is given by those characteristics in which it may vary without thereby ceasing to be what it is (sc. a table or an inadvertent act). Thus, the shape of a table is part of its matter, since it may vary in shape and still remain a table. In contrast, the *form* of a table is given by 'an answer to the question why we call a large variety of objects tables, and refuse the word to other objects' (*ibid.*). Now on this definition it may seem to follow trivially that 'it is the formal element which makes a thing to be what it is' (p. 15), and the interesting question is how we determine what is to count as the formal element of any given notion. Kovesi suggests (p. 3) that it is our reason for having tables which constitutes the guiding principle for deciding what are and are not tables. More generally, the formal element of a notion is connected with the point of having that notion in our language (p. 15), and the point may be to identify, buy and sell something, or it may be to avoid or promote some type of behaviour (*ibid.*).

---

[1] Julius Kovesi, *Moral Notions*, London and New York, 1967. All page-references are to this work.

How clear is this so far? Take, for instance, the idea of a material element. It is true that a table can vary in respect of shape, colour, number of legs, etc., and still remain a table; but it is surely germane to point out that it must have *some* shape, *some* colour, *some* number of legs, etc., in order to be a table? Hence, although being square may be an inessential property of tables, having a shape is not—it is one of the properties that make a table a table. Again, why should it be thought impossible, as Kovesi contends, to derive the conclusion 'This is a table' from statements in the required range, such as 'This consists of a flat, mahogany surface resting on three iron legs'? Perhaps, as he suggests (p. 37), there will always be ingenious new ways of making tables quite different from the old, so that to that extent the concept of a table is "open-textured", but why shouldn't we be able to give *sufficient* conditions of this kind for an object's being a table? Kovesi's rejoinder is implicit in what has already been said. There might be a civilization which had objects physically rather like our tables, but which did not have our needs and social conventions. In that case, such a civilization would not have our concept of a table (*cf.* p. 5).

Such a rejoinder may be sustained, but it smacks of triviality. Either we make it a necessary condition for an object's being a table that it (usually) be used for the purposes which our tables are (usually) used for, or we do not. If we do not, there is no reason why such a civilization should not have our concept of a table and doubt is cast on what work can be done by 'the point of having tables' in giving the formal element of the notion. If we do make this a necessary condition then we may allow that the civilization in question does not have tables or the concept of a table; but someone may insist that at least the civilization has the concept *table**, instantiated by tables* which are physically indistinguishable from tables. And we may begin to wonder where such verbal manoeuvres are leading us.

However, it may be urged that such an unsympathetic attitude to Kovesi's rejoinder rests on a failure to grasp his main point, *viz.*, that very often we group phenomena together under some common name not because of any observable, empirical similarity amongst them but because of the point in using the common name in the first place (*cf.* pp. 16–17). You cannot, for example, *see* any similarity in different inadvertent acts, such as my knocking over a cup when I reach for the salt and my destroying a sandcastle when I jump back from the sea; you can only *see the point* of calling them 'inadvertent', *viz.*, to excuse them.

The difficulty here is that Kovesi does not make it at all clear what is to count as an empirical or observable similarity (Are the two terms used synonymously? Empirical or observable as opposed to what?). He does at one point draw a distinction between perceiving and knowing, and he says that 'we do not perceive something called murder' (p. 19), but with no further elaboration. In the absence of such elaboration, I see little reason to agree. Of course, perceiving two acts of murder may not be like perceiving two red balloons, but then Kovesi himself rightly castigates those who have taken colour-words as ideals to which all other descriptive terms must aspire (*cf.* p. 20). Two acts of murder will be similar by virtue of involving the deliberate putting to death of one human being by another who has no legal right to do so, etc. Doubtless this does not give sufficient conditions for an act of murder, and doubtless murder can be instantiated in a number of ways which are in other respects dissimilar (a stiletto instead of arsenic). But there still seems to be no reason to deny that such various instantiations are observably and empirically similar; it is just that the similarity is detectable at a higher level of abstraction than that of some other, simpler things. Likewise for two inadvertent acts: their similarity will consist in the performing of some action which it was not the agent's intention to perform—a similarity which is surely empirical?

It is unlikely that Kovesi would accept this suggestion since, as we have seen, he looks quite elsewhere for an account of why we call two phenomena 'acts of murder'. The further justification which he offers for so doing is slightly more familiar. It is that we have not been given the *meaning* of a term if we are informed only of the empirical similarities which recur in different uses of it. For suppose that the term 'tak' is coined explicitly to refer to geometrical figures with a pointed projection (p. 39). Kovesi objects that it would be wrong to say that we know the meaning of the term after the appropriate ostensive lesson: we need also to know the point of using the word, and how it might be needed in a way of life, otherwise we have not been given 'the rules for the proper use of the word' (pp. 39–40).

But it is absurd to say that this further information is necessary for an understanding of the meaning of 'tak' when such information might not exist at all. It might be that there is no further activity for the sake of which the given discrimination amongst geometrical figures is made: the point in discriminating and classifying tak-shaped figures may simply be to discriminate and classify them. If this meant that there might be notions which lacked a formal element then it would be so much the much the worse for Kovesi's thesis. However, we have not yet got to the bottom of the confusion involved in the idea of a formal element.

II

The extent of the criticisms made so far should not be exaggerated. If well-founded, the most they suggest is that Kovesi has expressed his thesis too generally. What he has to say, or part of it, might still be true of specifically *moral* notions. We may approach this by considering first another hypothetical example which he constructs (pp. 46 ff.). Bus conductors sometimes print a ticket of the wrong value for a given passenger, which they retain and give to some subsequent passenger. They may on occasion have to explain such conduct to an inspector. Kovesi then hypothesizes that the expression 'making a misticket' evolves in connection with such activities, first in the vocabulary of the conductors and then as an expression in common both to them and to the inspectors. Now we may concede that the expression *could* function as an "excuse-word" when it is common to both groups, as Kovesi suggests (p. 47), but it is quite wrong to say that 'the efficacy of the excuse is . . . admitted by the inspectors when the word is accepted in their common language' (p. 48). That depends entirely on what we choose to hypothesize about the history of the term. It is essential to recognize that all the options are open in this matter and we can hypothesize that both conductors and inspectors agree that mistickets are excusable, or that they agree that they are inexcusable, or that they reach no agreement at all, without in any way precluding the possibility that the notion of a 'misticket' should be formed. For the notion might be formed without excuses in mind at all but rather, say, with the intention of cutting a long story short.[1] Or else it may be formed with excuses in mind, but come to be adopted by people who think differently on that question—consider how people may disagree about the justifiability of suicide and yet still be operating with what is in all other respects recognizably the same notion.

However, let us now suppose that the term 'misticket' *is* coined to excuse those who go through the performance associated with it. Kovesi points out that such a term may be misused in two different ways, by falling foul either of its material or of its formal element. Thus, a conductor may describe a regular, *used* ticket as a misticket (infringement of the material element), or an inspector may *rebuke* a conductor for issuing a misticket (infringement of the formal element) (p. 59). In order to substantiate further his claims about the centrality of the formal element, he then suggests that whether it is the formal or the material element which is offended against we draw attention to such misuses ultimately by appeal to the formal element (pp. 59–60). For if someone, so to speak, points to the wrong material elements we may remind him that if anyone were free to call anything he liked an x he would defeat the point of using 'x' to refer to some particular kind of phenomenon (p. 59).

This may be thought to reinforce a suggestion made earlier on Kovesi's behalf, and to imply a corresponding limitation on my criticisms. It does seem to be a central fact about specifically moral terms that there is a particular point behind their use, e.g., that 'inadvertent' is used to excuse and 'murder' to condemn. And while it is possible to imagine things differently in the case of hypothetical examples, this may not be so easy in the case of some actual notion already firmly embedded in an existing conceptual scheme.

---

[1] Someone might say that in that case the notion must be one other than that of a 'misticket'. But this is analogous to the (trivial) move of distinguishing the concepts of 'table' and 'table*', described in Section I.

Let us consider this. Someone may say

> (a) I wonder if murder really is the deliberate killing of one human being by another, etc. (*or* Murder is not the deliberate killing . . .)

or

> (b) I wonder if murder really is wrong (*or* Murder is not really wrong).

I take it that these two statements correspond to the two ways in which someone may misuse a term, i.e., that someone who utters (a) has infringed the material element in the notion of murder, whereas someone who utters (b) has infringed the formal element. Now it is true, given things as they are, that we are likely to be puzzled on hearing someone seriously and sincerely utter (a) or (b). But it does not follow that our puzzlement will be of the same kind in each case, nor that it will ultimately have the same source, which is effectively what Kovesi claims. Fairly obviously, the first suspicion on hearing someone utter (a) would be that he was not speaking the same language. Such a suspicion can be tested by asking the utterer what he does think murder is; if we receive a reply sufficiently far removed from our own (e.g., 'Murder is promising to marry someone and then failing to do so') then we shall be able to conclude that we are in disagreement with the utterer over words and how they are to be used. But it would be quite wrong to characterize a disagreement with the utterer of (b) in the same way. Provided he uses the expression with reference to the appropriate circumstances, *viz.*, the deliberate killing of one human being by another, etc., then we should not be in the least puzzled about *what* he was saying, though we might well be puzzled about *why* he should want to say it. As a consequence, we might conclude that the utterer of (b) was mad, depraved or psychopathic, but we ought not to conclude that he was guilty of conceptual impropriety. If we do so, we run the risk of precluding or trivializing a genuine moral disagreement. For suppose we told him that he would defeat the point of having the concept of murder if he went around saying it was not wrong: he could then make the trivializing move already noticed, and conclude that in that case he must be operating with the concept of *murder**, which differs from the concept of *murder* only on that one, apparently crucial, point. Indeed, it might be said with some justice that such a move would have been forced upon him. But to characterize the situation in this way would be to misdescribe it; it would be grotesque to suppose that we first of all have to persuade the man to start speaking our language and only then could we bring out a substantive moral disagreement.

This connects with the idea of a formal element and finally enables us to see what is at the root of Kovesi's confusion. We may recall that the formal element of a term 'x' is said to be given by our answer to the question 'Why do we call some things x and not others?'. But significantly different *kinds* of answer are possible to such a question, suggesting that the question itself is ambiguous. One kind of answer—call it a *C-type* answer—is 'Because all the things we call x have something in common, e.g., they are all cases of a ticket printed in error to the wrong value'. Now if *C-type* answers are taken to give the formal element of a notion, it is the man who utters (a) and the man who describes an already-used ticket as a misticket who are infringing the respective formal elements. And insofar as the formal element is supposed to tell us 'what a thing is' it seems to me that a *C-type* answer is appropriate (though, as I have argued, this still seems to be a matter of pointing out *empirical* similarities). Yet it is precisely the man who describes a used ticket as a misticket that Kovesi gives as an example of someone's infringing the rules for the *material* element (p. 59). How is this to be explained?

The solution lies in another kind of answer we can give to the question 'Why do we call certain things cases of x?', where this is understood to ask after the *purpose* with which, or the point of view from which, we label certain phenomena 'x'. And here the answer—call it a *P-type* answer—might be 'To isolate the cause of slumps' or 'To excuse'. It is with the *P-type* answers in mind that Kovesi is able to give as an example of misuse of the formal element an inspector's *rebuking* a conductor for issuing a misticket. Thus it becomes apparent what is wrong with Kovesi's claim that in both types of misuse (formal and material) we appeal ultimately to the formal element. In the case involving rebuke, we appeal to the *P-type* formal element (the notion was formed in order to excuse); in the case involving the description of a used ticket as a misticket, we appeal to the *C-type* formal element (the notion was formed in relation to unused tickets). This does not show that there is anything in common between the two cases except the ambiguous expression 'formal element'.

It is, in part, this equivocation which vitiates Kovesi's further views on moral notions. It is a *C-type* account which gives the meaning or essence of a notion. Kovesi's remarks about moral notions are chiefly of the *P-type*, but the success of his argument depends on these remarks' having the centrality and priority of a *C-type* account, a centrality and priority which they do not have even in the case of *moral* notions.[1] For although it might be true that the point of having the notion of murder is to condemn, the point is to condemn *a certain kind of action* which can be discriminated from other kinds of action only by means of a *C-type* criterion. It is this priority of the *C-type* element which makes moral disagreement possible: someone may—what is essential—agree with us on what is to count as another instance of the same thing, but disagree with us with regard to condemning it.

---

[1] But see Section III for important qualifications to this claim and several others made in the remainder of Section II.

However, a later sophistication in Kovesi's position may be thought to deal with this difficulty. He distinguishes between *complete* and *incomplete* moral notions (p. 109). 'Killing' is a notion not formed completely from the moral point of view, since it ranges over acts some of which are morally right and some wrong. Consequently, we may use the terms 'right' and 'wrong' to discriminate, within this class, acts which are different from a moral point of view. Such a notion is contrasted with a notion like 'murder', in connection with which 'the words "right" and "wrong" are used only as reminders, they remind us what was the point of forming such notions' (*ibid.*). Kovesi never actually says that 'Murder is wrong' is a tautology, but I take it this comes close to expressing his meaning.[1]

But this raises the question whether we are in fact possessed of any such complete moral notions.[2] Take, for instance, the notion of *lying*. Kovesi holds that it is an incomplete moral notion because some particular acts of lying may not be wrong—witness the familiar case of the homicidal maniac asking the whereabouts of his chosen victim. However, he suggests (p. 108) that such a notion could *become* a complete moral one if we had a special term for such cases, and he coins the term 'savingdeceit', whose *C-type* element is 'saving a life by means of some act of deception'. Then, he says, 'we should be more willing to say that lying is always wrong, or simply that lying *is* wrong. We would also now have a new "principle": "savingdeceit is right (or good)", and in situations where we would perform an act of savingdeceit we would no longer be confronted by a "conflict of principles"' (*ibid.*). There is, perhaps, an implicit suggestion here that once we have the term 'savingdeceit' then my behaviour in the would-be homicide case can no longer be characterized as lying. But this would be incorrect. It is true that acts of savingdeceit will not *necessarily* be acts of lying—there are other ways of saving a life by means of an act of deception than saying what is false. But lying is still one such way and it is the way involved in our example, which thus provides an instantiation both of lying and of savingdeceit.[3] Consequently, the 'conflict of principles' may not be so easily resolved, a point which will receive further attention in a moment. Moreover, even if the action in the example did cease to be an act of lying, I should not feel any more disposed to accept the suggestion that lying is always wrong, *if this is equivalent to the claim that it is always wrong to tell a lie.* It is not wrong to tell a lie for a harmless joke; in certain circumstances it may not be wrong to tell a lie in order to avert a great deal of suffering, or to boost someone's morale. And such a list is indefinitely extendable. Now, of course, we could coin new terms to express the *C-type* element in each of these cases, we could talk of 'jokingdeceit', 'boostingdeceit', etc. But even if (as is false) an action falling under one of these terms could not also be an act of lying, the very open-endedness of the list suggests that if we are to save the thesis that it is always wrong to lie then we shall need to coin some general term whose *C-type* element is something like 'saying what is false with intent to deceive, where doing so is morally acceptable'. And no one is likely to disagree with the claim that it is wrong to tell a lie except when it is not.

---

[1] He says, for example, 'If someone understands the notion of murder or stealing, to say that they are wrong does not give him any more information' (p. 26).

[2] A question which is not rhetorical. I believe that in a sense we are, but not in the sense Kovesi intends. See, again, Sections III and IV.

[3] Kovesi appears to concede this on pp. 128–9.

These comments are made in connection with lying, according to Kovesi an incomplete moral notion which can be made complete. They apply equally to allegedly already-complete moral notions, such as stealing and murder. I believe, as a matter of moral opinion, that it is sometimes not wrong to steal, and I reject the suggestion that I commit any linguistic or conceptual error when I say as much. It is perfectly possible for someone to say the same in connection with murder.[1] Rather than refusing to understand such a man, it would be slightly preferable to refuse to talk to him on the grounds that he showed a corrupt mind. But better still to do neither: better to argue with him instead.

II[2]

'But surely saying that murder is wrong is not just like saying that clouds are grey? Surely there is something special in the nature of the claim, which deserves to be brought out?' Such a complaint is justified. In this and the following section I shall attempt to meet it and at the same time indicate how the issues discussed relate to the problems of moral decision-making.

As well as expressing far more general moral notions than terms like 'courageous', 'dishonest', 'treacherous', the terms 'right' and 'wrong' sometimes carry a specially decisive force when used in particular situations.[3] So, when I say that murdering John was wrong I may be drawing attention to some general characteristic of the action, saying that it exemplified a type of action which is wrong, or else I may be suggesting that it was the wrong *thing to do*. In the light of this contrast between different uses of 'wrong' it is possible (and, as I shall try to show, profitable) to distinguish between the predicates $wrong_G$ (i.e., wrong in some general characteristic) and $wrong_D$ (i.e., wrong "decisively"). We may then relate this distinction to—but *not* identify it with—the distinction between action-types and actual, concrete actions. We sometimes wish to talk of actions of some particular *kind*, and we do so by means of a concatenation of general descriptive expressions—'murder', 'catching flies', 'putting a friend to bed', etc. We also sometimes wish to talk of particular, historical actions, and this we do by adding to an appropriate description some "individuator" which locates the action as a concrete, historical occurrence—'the murder of John F. Kennedy', 'putting John to bed last night', etc. Paradigmatically, such individuators will involve reference to some spatio-temporal location, since this is the most reliable way of securing unique reference. The description of an action-type may, of course, be highly specific, provided only that the description is in principle instantiable more than once. Thus, 'the murder by shooting of a Catholic President travelling in a motor car' is a description of an action-type rather than an action-token.

---

[1] Professor Strawson has pointed out to me that Auden found it convenient at one time to speak of 'the necessary murder'.

[2] The arguments in this section have benefited from discussion with Adrian Ashley.

[3] In this respect they may be bracketed with 'ought'. *Cf.* Bernard Williams, 'Ethical Consistency', *Supplementary Proceedings of the Aristotelian Society*, 1965, pp. 123–4.

Now $wrongness_G$ is primarily applicable to action-types. This is obvious enough when it is a predicate whose point is to attach to certain general characteristics and when action-types are themselves entirely specified by such characteristics. But $wrongness_G$ will also be derivatively applicable to particular actions. Any actual action, in all its concrete particularity, is an instantiation of indefinitely many action-types; hence, if murder is $wrong_G$ and I perform the act of murdering someone, then that act is, so far, $wrong_G$. The qualification 'so far' is inserted here because $wrongness_G$ is applicable to particular actions only 'under a given description'. For suppose that my concrete action (call it 'A') can be truly described as 'stealing' and as 'saving a life' and that stealing is $wrong_G$ but saving a life is not $wrong_G$. Then it might seem that we could make the following inferences

(a1) Stealing is $wrong_G$
(a2) A is stealing
(a3) Therefore, A is $wrong_G$

(a4) Saving a life is not $wrong_G$
(a5) A is saving a life
(a6) Therefore, A is not $wrong_G$.

Hence, it may seem that we are committed, by virtue of (a3) and (a6), to ascribing contradictory predicates to A. But the ascription of a G-subscript predicate to a particular action is indissolubly linked to that action's having a certain general character, and in a way does not go beyond this. Compare how my contribution to a debate may be helpful, insofar as it soothes frayed tempers, but not helpful insofar as it contributes nothing new to an understanding of the issues. Here again, contrary to appearances, there is no ascription of contradictory predicates. Hence, statements like (a3) and (a6) must be qualified: A is wrong_G *qua act of stealing*. With that proviso there is nothing wrong with these inferences, and if I hold that a certain *kind* of action is wrong_G then I do become committed to the wrongness_G of a *particular* action of that kind. Indeed, there is reason to say that the ascription of any predicate to an action-type is always to be explained and unpacked, at least in part, by pointing out that this entails a similar ascription to any particular actions of that type.

In contrast, D-subscript predicates attach primarily to particular actions rather than action-types. This is again a reflection of the reason why we need such predicates at all. There are times when a decision has to be taken about what to do, when it is not enough simply to list the pro's and con's of different courses of action but also necessary to opt for one. Such a choice will be guided, at least for an agent who is trying to act morally, by determining what D-subscript predicates attach to the different concrete courses of action, for it is particular actions in all their concrete specificity which are actually performed.

Now I should like to be able to go further, and say that D-subscript predicates are *only* applicable to particular actions. My reason would be that, in practice, however complex one makes the specification of an action-type it will not be possible to rule decisively for or against it in the way implied by attaching to it a D-subscript predicate. Because of the commitment with regard to particular actions which such an ascription would entail, and because we cannot know in advance in just what circumstances a particular action answering to the given type will turn up, we ought not to stick our necks out so far. But apart from the fact that my reasoning here might be thought to embody a mistake (which I consider in the next paragraph) it would be wrong to represent the inapplicability of D-subscript predicates to action-types as a matter of necessity. The constraints here are empirical and moral rather than conceptual, and one ought to resist the standing temptation in philosophy of representing what is desirable as somehow inescapable. If someone[1] really thinks, in advance, that an action-type specified in entirely general terms can safely be pronounced the wrong (or right) thing to do without qualification, he is entitled to his opinion—but I should certainly want to argue with him.

But does the commitment running from action-types to particular actions really provide a reason for withholding D-subscript predicates

[1] E.g., Professor Anscombe. *Cf.* 'Modern Moral Philosophy', *Philosophy*, 1958.

from action-types? If we can hold with regard to G-subscript predicates that the inference from type to concrete action is valid only in a referentially opaque way, why cannot we say the same thing here? If we can, then the commitment in question is far less onerous: all that will follow from the claim that φing as a type is wrong$_D$, with regard to any concrete action, is that that action is wrong$_D$ *qua act of φing*. My unsatisfactorily brief reply to this suggestion is that if you treat a D-subscript predicate in this way it is no longer a D-subscript predicate—it has lost the decisiveness which is the core of its meaning. Such a predicate is formed for application to an action, not in a two-dimensional way under one of its aspects but in its entirety, for it is actions in their entirety that we have to live with in a moral context. An agent with a desire to act morally will attempt to look at each course of action in this way, in its totality. If he does not do this then he is not attempting to determine whether a D-subscript predicate applies to it.

Not that anyone is forced to do so in any case. My proposed division of predicates is not forced upon us *a priori*; it is one which the contingent complexity of human life makes advisable. To that extent we have an option in adopting the proposed schema, just as we have an option in deciding whether each type of predicate is attachable both to action-types and to particular actions. However, *if* we choose to adopt the schema then we do become subject to certain constraints as a matter of necessity, such as that D-subscript predicates attach to a particular action under any true description. The most important of such constraints, which I wish to insist upon, is that *one cannot infer the wrongness*$_D$ *of a particular action from its wrongness*$_G$. In a particular situation it may be that the only courses open to me are φing and ψing, each of which is wrong$_G$. But it will not follow that, say, φing is also wrong$_D$, the wrong thing to do. On the contrary, if the only available alternative is far worse than φing then it may actually be wrong$_D$ *not* to φ. Alternatively, take the case where one course of action is wrong$_G$ under a certain description and not wrong$_G$ under another. If the kind of inference in question were valid, then we should have to conclude that the action was the wrong thing to do and not the wrong thing to do. In other words we should once again have lost the unequivocal, decisive verdict on courses of action which it is the role of D-subscript predicates to express. Hence it is a necessary condition of accepting *this* schema that such an inference be disallowed.

The pressure to accept the schema comes from the connected facts that human actions have many effects in many different directions and that they can be truly described in an indefinite number of different ways. In consequence, moral situations are generally infinitely more complex than the standard text-book examples suggest. Kovesi himself draws attention to this (*cf.* pp. 112, ff.) and yet he fails sufficiently to appreciate it. For example, in the inevitable discussion of Sartre's student he implies that the answer to the dilemma is given by whether the act of joining the Resistance constitutes leaving, forsaking, abandoning or deserting a mother, and if none of these, what other act it constitutes (p. 117). It is equally disconcerting to find him talking of 'the' morally relevant description of an action (p. 110). But consider the number of factors which would be relevant to the student's decision. What sort of person is he? How well does he understand his own motives? Is there any possibility that in his heart of hearts he has already decided what to do? Is either choice such that he would be incapable of implementing it with determination? What sort of relationship does he have with his mother? What sort of woman is she? What are her views on the matter? etc. It is not that such information would give an automatic answer to the dilemma, but rather that an adequate answer is impossible without it. For each alternative course of action there will be not one but a plurality of morally relevant descriptions; and it is likely that for any one alternative course some true descriptions will carry an implication of wrongness$_G$ and others an implication of not-wrongness$_G$. Far from "the correct" description providing an answer to the dilemma, we have somehow to go beyond the stage of collecting such descriptions and achieve some kind of summation before we can reach (or more likely invent) a compendious description sufficiently comprehensive to indicate all aspects which are morally relevant.

What is at issue here is the connection between one's particular moral decisions and one's more general moral views, and I suspect that at this juncture Kovesi falls prey to one part of the orthodoxy which he is in the main concerned to refute. It may well be that the nature of moral reasoning (or of *any* reasoning) is such that the particular decision I reach will carry implications beyond the particular case, that my decision will generate principles, universalized maxims or whatever. But it is grotesque to suppose that this will be of any use to me in reaching a particular moral decision, to suppose that I could be certain of something of more general scope and at the same time in doubt about some particular instantiation of it. It is putting a similar cart before a similar horse to suppose that as a first step I can arrive at *the* correct, morally adequate description of what I propose to do, and that this will enable me to take the next step of deciding whether it is what I ought to do.

IV

By way of summary and conclusion, I want to offer a model of the moves which are typically involved in moral decision-making and indicate where the difficulties lie in making such moves. Two points should be borne in mind. First, it is not suggested that any actual piece of moral reasoning will display such a neat, linear form, only that it will approximate to it. Secondly, the model is offered with a serious and sincerely committed moral agent in mind, i.e., a man who wants to make the required moves and arrive at a moral conclusion. It may not be appropriate, therefore, for a philosopher who is plumbing the outer limits of intelligibility by seeing what he can "deny without self-contradiction".

A preliminary distinction is needed between *primary* and *secondary* moral concepts and terms What I have in mind is the distinction between terms like 'right', 'good', 'wrong' (primary), and 'dishonest', 'brave', 'generous', 'sympathetic' (secondary). There are at least two grounds for the distinction: secondary terms are far more concrete descriptive terms, in that what is to count as, e.g., a generous act is fairly clear and definite in comparison with what is to count as a wrong act; and whereas in the case of at least some secondary terms (e.g. 'humble') it is possible to dispute whether they carry negative or positive moral force, this is not so in the case of primary terms. It may seem that all I am doing here is reproducing the orthodox descriptive/evaluative dichotomy. I do not believe that is so, but even if it were it is of the first importance that the distinction here is made *within* the class of moral concepts, not between them and some other class. On the other hand, it may be that further scrutiny of the primary/secondary distinction would reveal not a dichotomy but a continuum of concepts, of greater and lesser degrees of descriptive specificity and moral explicitness (e.g., is 'just' a primary or a secondary moral term?). If so, I suspect that many "descriptive" terms could be annexed to the end of such a continuum without looking at all out of place. The arguments of Professor Anscombe's paper 'On Brute Facts'[1] are relevant here, and deserve more elaboration than anyone has yet given them.

The moves which I now suggest are typically involved in reaching a moral decision are these

(1) *The move from "factual" considerations to conclusions involving secondary moral concepts.* E.g., the move from the fact that X killed Y to the conclusion that X murdered Y, or from the fact X ran away to the conclusion that X behaved in a cowardly way

(2) *The move from secondary moral considerations to conclusions involving primary moral concepts in their G-subscript use.* E.g., the move from the fact that X behaved in a cowardly way to the conclusion that in so doing he acted wrongly$_G$

(3) *The move from primary moral considerations in their G-subscript use to a conclusion involving a moral concept in its D-subscript use.* E.g., the move from the fact that X's action was wrong$_G$ to the conclusion that it was *the wrong thing to do.*

---

[1] G. E. M. Anscombe, 'On Brute Facts', ANALYSIS 18.3, 1958.

The orthodoxy which Kovesi opposes makes much of the supposed difficulties involved in the first move. Effectively, Kovesi makes no less of them; he merely insists that there are difficulties as great for many moves from "fact" to "fact". Whilst it would be an exaggeration to say that there are *no* difficulties here, I believe that they are not serious in either context.[1]

In practice there is often no difficulty in making the second move. That is the degree of truth in Kovesi's contentions, though it becomes distorted if he compresses into one (as I think he does) the second and third moves. To do so is to slur over the crucial difficulties of making the third move, the move to a definitive moral conclusion.

The third move is problematic for practical and theoretical reasons, but it is a move which has to be made. It can be avoided, to be sure, in retrospective moral thinking where an unequivocal final verdict may not be needed. But for moral *agents,* who have to choose between alternative actions, it is inescapable. For reasons already sufficiently stressed, it is impossible for such an agent to confine himself to secondary moral concepts and discard D-subscript predicates on the grounds that they carry only "mesmeric force".[2] The practical difficulties arise from the complexity of human action and interaction. Quite simply, it is not easy to be certain that one has collected all the secondary moral descriptions of alternative courses of action which might influence one's conclusion. But the theoretical difficulty is how one is to arrive at that conclusion even if one does have the necessary totality of secondary descriptions. When I am confronted with different possible courses of action each of which is morally characterizable in different ways, so that I have moral reasons both for performing and for not performing each of the alternatives, how do I decide what to do? According to what principles am I to effect the required summation? This is not, of course, a problem for someone with a unique moral principle, such as an act-utilitarian; but it is a problem for anyone who attaches any kind of intrinsic importance to secondary moral descriptions, and there are good reasons for doing so.[3] It is at this juncture that we have a kind of freedom, whereas Kovesi seems to imply that having a particular set of "complete" moral concepts will answer the question for us. 'Freedom', however, is probably a misnomer for such a burden, for what we are confronted with is the whole problem of balancing competing moral considerations in a rational way.

---

[1] For some reasons, see Anscombe, *ibid.*

[2] As I understand it, this is Professor Anscombe's suggestion in 'Modern Moral Philosophy', *loc. cit.*

[3] *Cf.* Williams' suggestive remarks on the rationality of regret, *op. cit.*

University of Bristol © KEITH GRAHAM 1975

INDIRECT REFERENCE

By THOMAS BALDWIN

IN his famous paper 'Ueber Sinn und Bedeutung' Frege argued that a referring expression does not have the same reference when it occurs within indirect speech as it has when it occurs within direct speech (i.e., as I shall say, its indirect reference and its direct reference are not the same). Frege's arguments will be familiar, and relate to the fact that ordinary identity statements do not licence substitutivity within indirect speech. Frege further argued that, since referring expressions with the same sense can be substituted *salva veritate* within indirect speech, one can identify the indirect reference of a referring expression with its sense.

Frege's arguments are powerful and compelling; but in this paper I shall initially discuss their application to sentences in which indirect speech verbs are iterated, a matter which, so far as I know, Frege never explicitly discussed. Frege did say, however, that when a sentence occurs within indirect speech, it does not have the same sense as when it occurs in direct speech (i.e., as I shall say, its indirect sense is not the same as its direct sense). Since, for Frege, sentences are referring expressions, this suggests that, in general, the indirect sense of a referring expression is not to be the same as its direct sense. And this, in turn, suggests that the doubly indirect reference of a referring expression is not to be the same as its singly indirect reference (where the doubly indirect reference of an expression is its reference when it occurs within the scope of two indirect speech verbs, etc.). For, given that singly indirect reference is the same as direct sense, it is natural to suppose that doubly indirect reference should be the same as singly indirect sense.

If one extrapolates Frege's ideas in this way, one will be led to postulate a hierarchy of indirect senses and referents, each one induced by the addition of a further indirect speech verb. This is certainly the sort of theory which Church discusses as a development of Frege's ideas in his paper 'A formulation of the Logic of Sense and Denotation' (pp. 3–24 in *Structure, Meaning, and Method*, ed. Henle, Kallen, & Langer, New York, 1951). But in his recent book (*Frege: Philosophy of Language*, London, 1973, esp. ch. 9), Dummett argues that such a theory should be rejected. Instead, Dummett maintains, indirect and direct sense should be identified, and, in this way, through the identification of doubly indirect reference with singly indirect sense, reference will be limited to two levels, direct and indirect. Dummett has two arguments for this view. His first argument (op. cit., p. 267) is that unless one identifies indirect and direct sense, one cannot say what the indirect sense of a referring expression is, and hence, within the first Fregean theory, one cannot give a full account of the truth-conditions of sentences with more than one indirect speech verb. This argument clearly has considerable force, which is further reinforced by a second consideration. Dummett suggests that behind Frege's distinction between direct and indirect sense lies the argument that, if the reference of an expression is a function just of its sense, then, given that indirect reference differs from direct reference, indirect sense must differ from direct sense. Dummett argues that the initial premiss of this argument is mistaken: the reference of an expression is to be regarded as a function not only of its sense, but also of the context in which it occurs. In this way one can allow for a difference between direct and indirect reference without the need for any corresponding difference between direct and indirect sense.

Dummett's objections to the first Fregean theory seem soundly based; but is Dummett's own simplified theory free of trouble itself? Dummett retains Frege's identification of indirect reference with sense, and, in the context of his theory, this requires Dummett to hold that identity of sense is always sufficient for substitutivity within indirect speech, however many times indirect speech verbs are iterated. Yet such a view can be made to appear implausible. A good example is suggested by Prior (cf. *Objects of Thought*, Oxford, 1971, pp. 54–56): the sentences

(1) X wonders whether the proposition that Jones wonders whether all unmarried men are unmarried is the very same proposition as the proposition that Jones wonders whether all unmarried men are unmarried

and

(2) X wonders whether the proposition that Jones wonders whether all bachelors are unmarried is the very same proposition as the proposition that Jones wonders whether all unmarried men are unmarried

do not look equivalent; for (1) makes X look crazy, but (2) only makes him look like a very cautious philosopher.

Prior holds, as Dummett must, that in fact (1) and (2) are equivalent. Prior argues that any impressions to the contrary are the result of confusing (1) and (2) with

(3) X wonders whether 'Jones wonders whether all unmarried men are unmarried' and 'Jones wonders whether all unmarried men are unmarried' express the same proposition

and

(4) X wonders whether 'Jones wonders whether all bachelors are unmarried' and 'Jones wonders whether all unmarried men are unmarried' express the same proposition

or with some similar pair of sentences about sentences which are used in (1) and (2). Now it is certainly clear that (1) and (2) are distinct from, but easily confused with, (3) and (4). The trouble with Prior's argument, however, is that in one situation there is an easy path from the non-equivalence of (3) and (4) to the non-equivalence of (1) and (2). For suppose the following is true:

> (5) X knows that 'Jones wonders whether all unmarried men are unmarried' unambiguously expresses the proposition that Jones wonders whether all unmarried men are unmarried and that 'Jones wonders whether all bachelors are unmarried' unambiguously expresses the proposition that Jones wonders whether all bachelors are unmarried.

If one assumes, as seems reasonable, that if a rational person knows that $a=b$ and wonders whether Fa, then he wonders whether Fb, then it follows, given (5) and X's rationality, that if (1) and (2) are equivalent, so are (3) and (4); and hence that if (3) and (4) are not equivalent, neither are (1) and (2). As a result, the introduction of (3) and (4), far from making more secure the equivalence of (1) and (2), gives one a way of arguing that they are not equivalent.

In order to preserve the equivalence of (1) and (2), therefore, Dummett is committed to the view that if (5) obtains, then, given X's rationality, (3) and (4) will be equivalent; and thus, in general, to the view that if any two expressions have the same sense, then, where a man knows the sense of each of them, he knows that they have the same sense. This is, indeed, a view which Dummett is happy to maintain—he calls it a 'compelling principle' (op. cit., p. 95). Dummett does not, I think, offer any argument for this view, and although it is a crucial element in his position, it is not obvious that it is correct. It seems possible to envisage a situation in which grammarians who, by any ordinary standards, know the meaning of a pair of synonymous expressions, get themselves confused when they are theorizing about these expressions and come to think that the expressions are not synonymous. Indeed, I would be inclined to recognize such a situation in the passing fashion for attention to presuppositions in grammatical theory. There are those who have thought that sentences such as 'Mary loves John' and 'It is Mary who loves John' differ in sense, because the latter presupposes that someone loves John in a way in which the former does not. I do not myself incline to this view; but in taking these two sentences to have the same sense, I do not feel obliged to suppose that those who have thought otherwise have not properly grasped the sense of these sentences. Rather, I merely suppose that they have confused pragmatic with semantic features of these sentences.

Knowledge of the sense of an expression is not an all-or-nothing business, but, like most skills, admits of degrees, and it would be silly to deny that someone who recognizes that two synonymous expressions are synonymous displays a better grasp of their sense than someone who fails to recognize it. I want to suggest, however, that the situation with regard to mistakes about the identity of the senses of expressions is no different from that which arises quite generally with regard to mistakes about the identity of abstract objects. Suppose a child knows perfectly well how to operate with fractions and with their decimal expansions, but (like most of us) sometimes gets confused about the relation between the two. It seems to me that one would normally think that where the mistake is a complex one this does not show that the child does not know what the number in question is, although where the mistake is a very simple one it may show just this. There is, I suggest, a natural range of reactions to mistakes about the identity of abstract objects. Where the object is a relatively simple one a mistake about its identity probably shows that one does not know what the object is; where the object is a relatively complex one there is no need to suppose this. Of course, in the case of numerical identities there are procedures for showing how questions about the identity of relatively complex numbers can be resolved into questions about the identity of simpler numbers; as a result anyone who persists with a mistake about the identity of a number in the face of such procedures can be judged, in the end, not to know what the number in question is. But the existence of such procedures does not show that, in the first instance, one must suppose that anyone who makes a mistake about the identity of a number thereby shows that he does not know what the number is. The situation is analogous to that discussed by Dummett with regard to the distinction between analytic truths and trivial truths (op. cit., pp. 288–90). Analytic truths follow by obvious steps from trivial truths: but this does not make them into trivial truths themselves.

In so far, therefore, as it is allowed that one can make genuine mistakes about the identity of abstract objects, it should be allowed that one can make genuine mistakes about the identity of propositions and concepts, and thus about the identity of the senses of linguistic expressions. The question that can now be raised is how the possibility of such mistakes can be fitted in with Dummett's theory of indirect reference, or indeed with any theory of indirect reference. It is clear, first, that it will not fit in Dummett's theory at all, since there is no way for Dummett to stop inferences from premisses of the form

X knows that the proposition that $p =$ the proposition that p
The proposition that $p =$ the proposition that q

to conclusions of the form

X knows that the proposition that $p =$ the proposition that q.

If one wants to stop this kind of inference within a theory that uses reference to abstract objects in the semantic analysis of indirect speech, one is naturally led to postulate a hierarchy of abstract objects as indirect referents of referring expressions within successively more complex indirect speech constructions (cf. Church, review of papers by Black and White, *Jnl. of Symbolic Logic,* 1946, pp. 132–3). The difficulty, however, for any such theory is that it runs foul of Dummett's objections to the original Fregean theory. In particular, there must be some procedure for specifying the indirect reference of a referring expression when it occurs within an indirect speech context of arbitrary complexity in such a way that the specification of the indirect referent matches in number of subordinate clauses the context in which the referring expression occurs. It is hard to see what such a procedure could be. The only obvious way of iterating subordinate clauses is to add successive occurrences of 'the proposition that . . . is true', but the equivalence of all propositions of the form 'the proposition that *p* and 'the proposition that the proposition that *p* is true' indicates that this kind of iteration is useless in the present context.

It seems to me, therefore, that if one retains the natural view that one can be genuinely mistaken about the identity of abstract objects of any kind, one must abandon the kind of reference to abstract objects that is an essential feature of Frege's idea that referring expressions have a different reference when they occur in indirect speech from that which they have when they occur in direct speech. It is not easy to say just what should be put in the place of Frege's idea; but one suggestion might be that where a referring expression R occurs within the scope of an indirect speech construction, one should determine the truth-value of any resulting sentence by considering, not what R does refer to, but what R is known to refer to, or some similar condition on R's reference.

University of York

© Thomas Baldwin 1975

CONDITIONALS, IMPOSSIBILITIES AND MATERIAL IMPLICATIONS

By PETER DOWNING

RECENT and fairly recent papers in ANALYSIS have contained arguments for and against the view that conditionals are non-truth-functional. I shall comment on some of these arguments and argue that conditionals differ not only from material implications but also from other propositions with which they can be confused.

My first point is that the contrariety of opposite conditionals, that is, the contrariety of conditionals If p then q and If p then $\sim q$, or If p then q and If p then r, where q and r are contraries, is the basis of many of our ordinary beliefs and judgments, beliefs and judgments which guide us—which affect out conduct. Consequently, and perhaps this will not be disputed, and certainly *should* not be disputed, the onus of proof is on anyone who rejects the contariety of opposite conditionals, not on someone who accepts it. For example, suppose Jones comes to believe that if he dies within the next week without having taken out a life insurance policy his family will suffer, and suppose someone, a philosopher, of course (!), knows of Jones's belief, uses what Dr. A. J. Dale, in ANALYSIS 34.3, quaintly describes as the *logician's* '$\supset$', and tells Jones that it is *also* true that if he dies within the next week without having taken out a life insurance policy his family will *not* suffer, the philosopher's comment (and not just because of annoyance at possible implications of what he has said) might well be met with derision, amusement, indifference or even anger and some consequent action, or perhaps the suggestion that the philosopher should concentrate on his own very strange business, and not try to interfere with practical matters. After all, Jones's belief might well be of considerable importance to him and influence his subsequent actions, but how is he to act on it appropriately if the philosopher's comment is also true?

I am inclined to think that in this area of logic the contrariety of opposite conditionals has a status similar to that of the contrariety of p and $\sim p$ in *general* logic, and that it is therefore inappropriate to try to give a satisfactory proof of it, a proof satisfactory in that its premisses are more obviously true than its conclusion. One can, of course, attempt persuasive arguments in favour of the principle. I have just attempted one, I shall give another below, and I gave one in *Mind*, Oct. 1961. Moreover arguments *against* the principle must be examined, and there is a case for drawing attention to confusions which might lead someone to reject the principle.

One of these confusions seems to have influenced Mr. Michael Clark in ANALYSIS 34.3. Thus on page 79 he mentions

If that is a perpetual motion machine, it will dissipate energy

If that is a perpetual motion machine, it will not dissipate energy

and suggests that respective utterers of these conditionals may rightly come to agree with each other's claim and conclude that perpetual motion machines are physically impossible.

This seems to me to provide an example of a confusion to which I drew attention in *Mind*, Oct. 1961, and also in *Proceedings of the Aristotelian Society*, 1958–59, namely that it is a mistake to think that a conditional If p then q is equivalent to or deducible from the physical impossibility of $p . \sim q$ [$I(p . \sim q)$]. This is because $I(p)$ strictly implies $I(p . \sim q)$, but $I(p)$ does not strictly imply If p then q, since from $I(p)$ nothing follows as to what will occur or be the case *if* p is true. Consequently $I(p . \sim q)$ is not equivalent to and does not strictly imply If p then q. But Clark's example is surely an instance of the general truth that, whilst $I(p . \sim q)$ is consistent with $I(p . q)$, $I(p)$ follows from their conjunction. Thus the sentences mentioned by Clark do not express conditionals although the word 'if' is used. One may be tempted, as I pointed out in my earlier papers, to make a deduction from $I(p . \sim q)$ using 'if'. Often, for practical purposes, this doesn't matter, but it must be understood that the conclusion deduced is not a conditional. Of course, $I(p . \sim q)$ is also non-truth-functional, but it is not a conditional.

In this connection it is perhaps worth pointing out that there seems to be no good case for using physical impossibility, even with added provisos, in an account of conditionals. Take, for example, a proposition of the form: p might be true but only if q, which can be taken to be of the form $I(p . \sim q) . \sim I(p)$. That this is not a conditional cannot be proved by the argument in the previous paragraph, but what can be pointed out is that when someone believes, judges, etc., that if p then q he does not *thereby* believe, judge, etc., anything entailing $\sim I(p)$, or for that matter anything entailing $\sim I(\sim p)$. He is only concerned with what will occur or be the case *if* p is true. Thus a conditional If p then q is, as it were, neutral as regards the physical possibility of either p or $\sim p$. This being the case, If p then q does not follow from $I(p . \sim q) . \sim I(p)$. The second conjunct of this is irrelevant to If p then q, and I have already pointed out that If p then q does not follow from the first conjunct. Once again, in a deduction from $I(p . \sim q) . \sim I(p)$ it may not matter if one uses 'if', but it should be understood that a conditional has not been deduced.

Clearly then an examination of the logic of conditionals should not be, and should not rely too heavily on, an examination of different uses of 'if'. Apart from the uses just mentioned, 'if' can be used to express material implications or logical connections; but, I shall argue, it is not then used to express conditionals. Perhaps the view that 'if' is not all that important is supported, slightly, by the fact that locutions which do not contain it can express conditionals, for example, 'Encouraged he will work well', 'Encouraged he would work well', or 'I had to jot that down or I would have forgotten it', though the last sentence, admittedly, expresses something in addition to a conditional. But surely, as I have implied at the beginning of this paper, with subsequent examples which support this view, it is better, in this context, to ask, not how we speak, but how we think.

This is relevant to other points made by Clark. For example in ANALYSIS 32.2 page 35 he suggests that if opposite conditionals were contraries every statement pair of those forms would be mutually incompatible. If this were a semantical point, i.e., that every such statement pair which express *conditionals* must be mutually incompatible it would be acceptable, but Clark's arguments suggest that he has only made a syntactical point. His example concerning a perpetual motion machine suggests this. Moreover, on the page just mentioned he considers:

(C) The match won't be cancelled

(D) Whenever it rains on the day of a match, that match is cancelled

(E) If it rains the match will be cancelled

(F) If it rains the match won't be cancelled

and suggests that (C) and (D) are compatible, (C) entails (F) and (D) entails (E), so (C) & (D) which is consistent entails (E) & (F). Therefore (E) & (F) is consistent.

(C) and (D) are no doubt compatible, and (E) and (F) are compatible if they are or express material implications, and 'if' can be used on this assumption. But the argument needs to be strengthened, for example, by showing that the sentences in (E) and (F) *can* only be used to express material implications. Moreover, someone who believes that conditionals are non-truth-functional can point out that if (D) is a general material implication it follows from

(G) It never rains on the day of a match

and that, given that (E) is a conditional, (G) does not entail (E), since from (G) nothing follows as to what will occur or be the case if (G) is false, i.e., if it does rain on the day of a match. All that follows from (G) is the falsity of the antecedent, or perhaps of what is implied by the antecedent of (E). Thus since (G) entails (D), but (G) does not entail (E), it follows that (D) does not entail (E).

Mr. John Young, in ANALYSIS 33.2, argues that (C) does not entail (F). Perhaps Young's most effective point is that it is possible to believe that the match will not be cancelled (perhaps on the ground that the weather forecast is favourable) and consistently refuse to accept (F). One answer, presented by Clark in ANALYSIS 34.3, was, in effect, to suggest that if one knows that the match is going to take place, one can infer that nothing, rain included, is going to prevent the match; so one can conclude that (F). But this answer is ineffective. Someone who believes that conditionals are non-truth-functional can simply point out that from the fact that the match will take place, or no doubt from (C), either of them being conjoined with (E), one can indeed infer that the antecedents of (E) and of (F) are false, but apart from (E) one cannot deduce what else will occur or be the case if the antecedents of (E) and of (F) are true. A fortiori one cannot make any such deduction from the fact that the match will not take place, nor from (C), if either is taken by itself, not conjoined with (E).

Another reply by Clark to Young's objection was to concede that if a man believes that the match won't be cancelled, but thinks that it would be if it were to rain, he is likely to be reluctant to assert (F) or to agree with someone else who asserts it; and Clark goes on to explain this in terms of conversational implications. However, although Young argues that someone who believes (C), can consistently *refuse to accept* (F), and this *perhaps* implies that (F) and the possible refusal are both expressed in language, this is not essential to the argument, since someone who believes (C) can consistently believe, or judge, that (F) is false.

In this connection another practical example is relevant. A surgeon may judge that if some physical condition C applies to his patient, the patient will survive [If C then S] and also judge that if $\sim$C then $\sim$S. The surgeon may then judge that S, on the ground that by operating he can ensure that C will apply to the patient. He may be aware of a medical theory which he is inclined to reject as being probably false and as having dangerous consequences if applied in certain situations. From this theory it follows, on this occasion, that if some other factor D applies to the patient, the patient will survive [If D then S]. The surgeon may momentarily consider this theory and its possible application to this situation, and consistently with judging that S, reject as false that if D then S. Indeed he may decide to operate in such a way as to ensure that C, but not D applies to his patient. Quite probably the surgeon may have to make quick decisions, swiftly decide that if C then S and if $\sim$C then $\sim$S, reject as false that if D then S, and decide to operate in order to ensure C but not D. He may well make these decisions without speaking, or perhaps speaking only to give instructions to his assistants. In such a situation obviously no question of conversational implications need arise.

This sort of example also supports the contrariety of opposite conditionals. If on this and similar occasions the surgeon, having been to some logic classes, judges that if C then S, but reflects that it may also be true that if C then $\sim$S, he will be in a very difficult situation, and will do well to retire from the medical profession.

On more than one occasion Clark makes distinctions such as that between 'if it rains' and 'if it were to rain'. Thus in ANALYSIS 34.3 when defending the view that (C) entails (F) he concedes that the person who deduces (F) may not be in a position to conclude that the match would not be cancelled if it were to rain, and in 32.2, page 35, he writes, 'No doubt there are events whose occurrence has not yet been made impossible, and if these *were* to happen the match would be cancelled'.

As a preliminary it can be noted that in the quoted sentence the implied possibility that there may be events whose occurrence *has not yet been made impossible* is irrelevant, since, as I have pointed out, a conditional If *p* then *q* is only concerned with what will occur or be the case *if p* is true, and does not entail, nor, it may be added, does the person who states it necessarily imply, that *p* is physically possible.

Reliance on distinctions such as that between 'if it rains' and 'if it were to rain' has been criticized, for example by Mr. J. S. Edwards, in ANALYSIS 34.3, page 90, on the ground that it confuses the proposition expressed with implications arising from the context of its utterance. Of course, 'were', especially if it is emphasized, can serve to imply the falsity, or doubt as to the truth, of the antecedent of a conditional, but this is additional to the conditional thought which is expressed. However, I shall use 'were', etc., in the following argument, and subsequently point out that this argument, whose main implicit premiss is the contrariety of opposite conditionals, not only proves a further difference between conditionals and material implications, but also provides some support for my view concerning this relatively unimportant linguistic matter. Consider the following

(a) If Jones were to leave at 6 o'clock he would not take a taxi and would miss his train

(b) If Jones were to leave at 6 o'clock he would miss his train

(c) If Jones were to leave at 6 o'clock and were to take a taxi he would not miss his train

(d) If Jones were to leave at 6 o'clock and were to take a taxi he would miss his train

(e) If Jones were to leave at 6 o'clock then whether or not he were to take a taxi he would miss his train.

(a) strictly implies (b). But (a) is consistent with (c). Therefore (b) is consistent with (c). Since (b) is consistent with (c), (b) does not strictly imply (d), because (c) and (d) are contraries. But (e) is equivalent to the conjunction of

(d) If Jones were to leave at 6 o'clock and were to take a taxi he would miss his train

with

(g) If Jones were to leave at 6 o'clock and were not to take a taxi he would miss his train.

Therefore, since (d) is a conjunct of a proposition equivalent to (e), a fortiori (b) does not strictly imply (e).

Thus (b) does not strictly imply (e), though it is surely obvious that (e) strictly implies (b). However, $((p.r) \supset q)\ .\ ((p\ .\ {\sim}r) \supset q)$ is *equivalent* to $p \supset q$. This provides a contrast between conditionals and material implications. Moreover a very basic difference is involved since (b) can, as it were, be 'strengthened' further by adding to the locution which expresses (e) phrases such as 'whether or not he hurried', 'whether or not he got a lift', and so on. This 'strengthenability' of conditionals is one source of the richness of their logic in contrast to the comparative poverty of material implication. It also supports my view of the linguistic matter considered above, since the fact that (e) strictly implies (b) but (b) does not strictly imply (e) can surely be stated, in conversation, by using indicatives and saying that

(e)′ If Jones leaves at 6 o'clock then whether or not he takes a taxi he will miss his train

strictly implies, but is not strictly implied by

(b)′ If Jones leaves at 6 o'clock he will miss his train.

The characteristics of conditionals implied by the above argument also emphasize, if any emphasis is necessary, their difference from logical connections. (b) above could be true just because (a) is true. Thus an ordinary 'unstrengthened' conditional can, as it were, rely on unspecified factors. For this reason (b) is consistent with (c). But (b) can be strengthened, as in (e), by ruling out such factors as irrelevant. There are surely no comparable features in the logic of strict implication or of entailment. If p strictly implies, or entails, q, this is strong enough "as it is". No question of relying on unspecified factors arises, there is nothing comparable with the consistency of (b) with (c), and adding 'whether or not' factors is irrelevant. The same applies, of course, to the principles used in mathematical proofs.

Conditionals, I suggest, their antecedents and their consequents, are concerned with 'occurrables', i.e. what can, logically, occur in the world of space and time. I shall not go into this matter in detail, but given this view it was not necessary for Young, in ANALYSIS 33.2, pages 59–60, to place a *restriction* on the contrariety of conditionals where the antecedent is self-contradictory, since such so-called conditionals are not conditionals. In this context Young was concerned with supposedly mathematical conditionals, but given the success of my argument in the previous paragraph, there are none such, and Young's suggestion that an alternative might be to construe the 'if's *of reductio proofs* as differing from both hooks and the 'if's used for conditionals, does not go far enough. If it is suggested that there may be conditionals with self-contradictory antecedents, but whose consequents are 'empirical', it would be fair, I think, to throw doubt upon this possibility by asking Jones, the Jones mentioned at the beginning of this paper, how he would operate with the 'conditional' that if he dies, etc., and also *does not die*, etc., his family will suffer, or to ask the surgeon whether his patient would survive if C applied and also *did not apply* to him.

The strengthenability of conditionals is relevant to a point made by Dale in ANALYSIS 34.3. Dale suggests that if conditionals are non-truth-functional there is no general method of proving that the principle of transitivity applies to them. This is correct. It can, I think, be successfully argued that certain *types* of conditionals are transitive, and that a principle somewhat similar to transitivity, namely that If p then q follows from the conjunction of If p then r with If $p.r$ then q, applies to 'ordinary' unstrengthened conditionals. However, from If A at t_1 then whatever happened in the meantime B at t_2, and If B at t_2 then whatever happened in the meantime C at t_3 (or If A at t_1 and B at t_2 then whatever happened in the meantime C at t_3), it does not follow that If A at t_1 then whatever happened in the meantime C at t_3, since the premisses, but not the supposed conclusion, are consistent with If A at t_1 and $\sim$B at t_2 then $\sim$C at t_3.

Another point made by Dale is that on the view that conditionals are non-truth-functional the law of contraposition cannot in general be shown to apply to them. I agree. Indeed, I argued in my 1959 paper, page 138, and in *Proceedings of the Aristotelian Society*, Supplementary Volume for 1970, and J. L. Mackie in *Truth, Probability and Paradox*, page 109, has argued more tentatively, that since If p then q and If p then $\sim q$ are contraries, whereas If p then q and If $\sim p$ then q are compatible, if the law of contraposition were applicable to conditionals, from this consistent pair it could be deduced that if $\sim q$ then $\sim p$ and if $\sim q$ then p which are contraries. Consequently contraposition applies to no conditionals, of whatever type.

It will be clear, I think, that from the contrariety of opposite conditionals, combined with other acceptable principles, it can be deduced that conditionals differ radically, and perhaps surprisingly, from material implications and from certain other propositions, but why, for example, should restrictions on transitivity, and the rejection of contraposition count, somehow, against the view that there is an area of logic, namely the logic of conditionals, which differs from more "orthodox", or more adequately explored, areas? No doubt radical versions of empiricism, such as those argued for by Ayer, involve rejecting the possibility of such an area of logic, and I suspect that empiricist views of one kind or another, in some cases perhaps only subconsciously held, often underlie this rejection. Conditionals, on my view, are of assistance in the solution of certain philosophical problems, and in particular can be used in an account of physical impossibility (not vice versa); but I cannot argue for this here. I can only suggest that conditionals are involved in much of our thinking, of our true beliefs and judgments, about the world in space and time, and that, subject to empiricist objections which of course must be taken account of, the logic of conditionals is *qua logic* of interest, and may well be more important than the logic of material implication. I have never been surprised by the fact that several students in my experience, especially during their first year while very open-minded, have expressed amazement that material implication should be regarded as so important!

University College London © PETER DOWNING 1975

FURTHER IN THE MODEST DEFENCE

By JOEL RUDINOW

OF the two criticisms of my paper 'On "the Slippery Slope"' (ANALYSIS, 34.5) offered by Ms. Lindsay (ANALYSIS, 35.1) the first does not interest me. I do not think I have mischaracterized the slippery slope argument; not at least by exaggerating or overstating its conclusion. As for Ms. Lindsay's claim that the slippery slope argument exposes the lack of any basis for a morally significant distinction between intra-uterine beings—that's just false.

The second criticism interests me more, but it is badly put as an objection against a defence of abortion founded on as incontrovertible a principle of morality as there is. This principle is, as far as I know, taken by all sides as a datum for the abortion debate: it is the principle that any restriction on individual liberty requires justification. On the other hand, Ms. Lindsay's remarks do suggest an alternative to the slippery slope argument which, if sound, might support a conservative position on abortion. This gives me an opportunity to pursue the defence of a more liberal position a step further. The argument might be called 'the argument from uncertainty'; it grants that the status of the foetus as a bearer of rights is controversial and argues *on that basis* to the conclusion that abortion is wrong except where necessary to save human lives. Here is one formulation.

(1) If the foetus has the right to life then to destroy the foetus is wrong except where necessary to save human lives.

(2) The status of the foetus as a bearer of the right to life is controversial.

∴ (3) It is not known that the foetus has not got the right to life.

∴ (4) To destroy the foetus is wrong except where necessary to save human lives.

This argument requires additional premisses if it is to go through, because all that is warranted by (1), (2), and (3) is

(4A) It is not known that to destroy the foetus is not wrong except where necessary to save human lives.[1]

Suppose the premiss is added that

(*) To destroy anything which is not known not to have the right to life is wrong except where necessary to save human lives.

Then the argument can go through. (4) follows from (3) and (*). But (*) is unavailable to the conservative since it is quite plainly false. There are numerous organisms which are not known not to have the right to life, but which may be destroyed for many reasons, not just when human lives are at stake. But I can think of only one other plausible way of getting from (1), (2) and (3) to (4) and that is by the addition of the premiss that

(**) To do anything which is not known not to be wrong except where necessary to save human lives is wrong except where necessary to save human lives.

Again the argument goes through. (4) follows from (4A) and (**). But (**) is question-begging.

---

[1] (4A) is interesting. I have my doubts about its truth. But as I have similar doubts as to the truth of (1) I do not doubt that (4A) is a consequence of (1), (2) and (3).

Dartmouth College

© Joel Rudinow 1975

D. Z. PHILLIPS AND 'THE INADEQUACY OF LANGUAGE'

By David A. Conway

(1) IN *Death and Immortality*, D. Z. Phillips argues that talk of 'surviving death' is nonsensical, meaningless talk. I do not intend here to attempt to evaluate the arguments he presents for this controversial, but quite possibly correct, claim. My present interest is in some remarks he makes in response to a hypothetical objection to his arguments.

> At this point a philosophical and popular agnosticism is likely to be introduced into the argument. Someone is likely to say that we just do not know what happens to a human being after he dies. Our language, it might be said, is too confined to tell us anything about the world beyond the grave.[1]

In reply to this, Phillips makes the remarks which I should like to examine. If the objection means

> that our language *as such* is inadequate to tell us anything about the world beyond the grave, the notion of inadequacy is being misused. Our language is not a poor alternative to other means of communication; it is what constitutes communication. To say 'We only have our language' in this context, is not like saying 'I only have English.' In the latter case one might say, 'If you could speak Welsh you'd see why *hwyl* is untranslatable.' But one cannot say, 'Because we only have our language we cannot say what the world beyond the grave is like.' There can be an inadequate use of language, but *it makes no sense to say that language itself is inadequate*.[2]

If correct, this would be of crucial importance to our understanding of the task of philosophical theology, for a large part of the history of that discipline can be construed as reactions to, attempts to cope with, the nearly universally accepted premiss that we, as finite, limited beings, cannot fully understand the nature of an infinite, unlimited being, as a result of which our language, as such, is inherently inadequate for expressing, and so for understanding, the nature of the deity. Thus, Phillips would dispose of much of philosophical theology, and, unless we are content with a "God" we can fully understand (a limited, perhaps Zeus-like, deity), with theistic religion itself. And he would do so in very short compass indeed. His contention, then, is, if correct, far more important than the brevity of its assertion would indicate. In this brief discussion I shall make some critical comments concerning the claim. Perhaps my attempts will lead Phillips or someone else inclined to this view to elucidate and defend it more seriously.

(2) The basic claim that 'our language is too confined' to allow us to know the answers to some important questions should be disentangled

---

[1] D. Z. Phillips, *Death and Immortality* (London: Macmillan, 1970), p. 14.

[2] Phillips, *op. cit.*, p. 15. Last italics added.

from some mistakes which are commonly made by its adherents. For instance, it will not do at all to try to back up the view that language is limited by saying exactly what it is that cannot be said. Thus, Frederick Ferré makes himself a possible subject of a sort of Kierkegaardian comedy when he disavows his ability to describe the weather on an unpleasant day: 'Perhaps the English language is not yet equipped to indicate the more-than-drizzling but less-than-sprinkling condition of the atmosphere.'[1] ("Tell me your name, little boy.' 'I can't; I can't say "William".') And Wittgenstein, in the *Tractatus*, left himself open to Ramsey's famous quip: 'What we can't say we can't say, and we can't whistle it either.'[2]

Nor will it do to say that since our language is confined, since there are things beyond our comprehension, we must not attempt to understand, but instead must simply believe. An exhortation to believe is relevant if in question are insufficiently evidenced propositions; if our puzzlement is about *what* is to be believed, such exhortations can only be misleading rhetoric.

In addition, the 'language is too confined' doctrine must not be taken as showing that contradictions are somehow "all right", as if contradictions are happenstance consequences of the limitations of our language, and that, if only it did not have those limitations, squares could be understood sometimes to have three sides and bachelors to be fine husbands. If Phillips is correct in thinking that 'survival after death' is self-contradictory, then he is also correct when he says that 'to say that our language is too confined to tell us anything about the world beyond the grave simply obscures the issue'.[3] If we have a contradiction, the notion of 'survival after death' is incoherent, and no doctrine about our language being limited changes that fact. But the fact that such spurious appendages to the view that our language is limited are untenable cannot be taken to show that our language is not limited. For those addenda *are* spurious appendages. The basic position may still be correct.

(3) The position that 'it makes no sense to say that language itself is inadequate' is, I think, intuitively, about as implausible a position as we can imagine a philosopher seriously maintaining. In fact, it seems that we can easily generate a *reductio* of it.

(a) It is not the case that language is inadequate is, if true at all, some sort of conceptual, or necessary, truth. For if it were a contingent truth, it would have to be established by comparing what can be said in language with "everything that is the case", including those purported things which are beyond our finite comprehension. And, in order to do this, we would have to be able to know the things which are beyond our finite comprehension, since only if we are capable of knowing such things can we find that there are none. Also, it is clear enough that Phillips himself considers the impossibility of language being inadequate to be a necessary impossibility ('*it makes no sense* to say that language itself is inadequate').

---

[1] Frederick Ferré, *Language, Logic and God* (New York: Harper, 1961), p. 153.

[2] Quoted in John Passmore, *A Hundred Years of Philosophy* (London: Duckworth, 1959), p. 364. I do not mean that one necessarily cannot indicate what *sort of* thing it is that cannot be said. Thus Ferré would have avoided the realm of the comic had he said, 'Perhaps the English language is not yet equipped to describe some conditions of the atmosphere', and one might disavow human ability to understand God by saying, 'Human language is not able to describe the nature of a being that is worthy of a worshipful attitude'.

[3] Phillips, *op. cit.*

Now, consider the fact that some words are untranslatable; Phillips gives the Welsh word '*hwyl*' as an example. Suppose that this word, or the entire Welsh language, did not exist. In that case, it would seem that language would be inadequate, since there would be something which cannot be said (i.e., what one says in using the word '*hwyl*').

> (b) The word '*hwyl*', or its equivalent, exists in some human language

expresses, then, a necessary condition for the truth of (a), and therefore (b) must itself be necessarily true if (a) is (since if (a) can be falsified by the falsity of a contingent proposition it cannot be necessary). But that the existence of a particular word (or its equivalent) should be necessary is an absurdity which I think few would be willing seriously to countenance.

As simple as the above argument is, we can easily enough give a simpler one for the same conclusion. Can we not imagine a very primitive time in which there was only one "language", one in which there are exactly three words or phrases, meaning, perhaps, 'Give me food', 'Give me sex' and 'Watch out for the diplodocus'? The possible existence of this communication system as the sole "language" is a sufficient condition for the falsification of the supposedly necessary truth, (a).[1]

These arguments depend on interpreting Phillips as saying that at any time the language that we actually have cannot be inadequate. Perhaps, however, he means instead to say that there cannot be things which cannot be expressed in any *possible* language. Thus, for example, no language may, at a given time, be capable of expressing what '*hwyl*' expresses, but it is clearly possible that a language should be able to do so.

This position is considerably more plausible, but it will not do to make Phillips' point. For it is compatible with the objector's point which is that our language is now too limited to express what it will be like beyond the grave. The objector can even say that our language, i.e., the language of mere earthly beings, will always be too limited to allow us to understand the future condition, that that condition is expressible in a language which we can learn and understand only when we are in that condition.

---

[1] Such a primitive signalling system no doubt should not be called a 'language' (hence the use of "scare quotes" above). Nonetheless, it would be 'what constitutes communication' and so, on Phillips' account, it cannot be inadequate.

(4) Does Phillips offer anything that could be construed as *argument* for his claim? If he does, it would appear to be contained in the following previously quoted sentences:

> If what the hypothetical objector means is that our language *as such* is inadequate to tell us anything about the world beyond the grave, the notion of inadequacy is being misused. Our language is not a poor alternative to other means of communication; it is what constitutes communication. . . . There can be an inadequate use of language, but it makes no sense to say that language itself is inadequate.

Apparently, this argument comes to this: to be inadequate is to fail to live up to some relevant standard. The purpose of language is communication, and so the relevant standard for adequacy is success at the task of communicating. But since language 'is what constitutes communication', it makes no sense to say it fails at the task of communication, since this would be saying that language fails to do what language does. Therefore, the claim that language is inadequate is incoherent.

There seems little reason why a person who believes that there are truths, specifically religious ones, which cannot be expressed in our language should be disturbed by this argument. For his position is that language is inadequate in the sense that with it we cannot express everything that is (or could be) the case. On Phillips' account it turns out that language cannot be inadequate, but he has made it turn out in this way by decree; he declares that language just is the standard of adequacy, i.e., language can express everything that can be expressed in language. This in no way meets the objector's point—there may be things which cannot be expressed in our language.[1]

(5) I said at the outset that Phillips' thesis, if correct, would be of ultimate importance for theism. That evaluation must now be amended; the thesis for which he apparently intends to argue, that there cannot be things which cannot be expressed in our language, would be of such

---

[1] The way in which I put this point makes it appear that it depends on there being some comprehensible notion of a "world" apart from our conceptual frameworks or "language games", an independent reality against which we can and must check our knowledge claims. The doctrine that there is such a reality is not acceptable to some neo-Wittgensteinians, including Phillips (see *The Concept of Prayer* (Routledge and Kegan Paul, 1965), p. 9), or to such philosophers as Kuhn, Quine, Sellars and Rorty. I shall restrict myself to two brief comments on this controversial position. First, in the argument which I am discussing, Phillips makes no mention of such a position; and so he apparently does not believe that his argument depends on that position. Second, even on the view that there is no knowable independent reality, 'there are things which cannot be expressed in our language' would seem to be perfectly sensible. To say this would mean that there are things which cannot be said given our present conceptual framework (or set of language games) but which could be said within other frameworks (specifically, within God's conceptual scheme).

great importance. The thesis which his argument actually supports, that language cannot be called 'inadequate' simply because it is 'what constitutes communication', is of little interest and bears not at all on whether language is adequate in the sense that everything that is the case can be expressed in a language. Phillips' apparently important thesis turns out to be either quite unsupported by his argument or to be a different thesis which is really not very interesting.

University of Missouri-St. Louis © DAVID A. CONWAY 1975

SUBSTITUTIONALISM AND SUBSTITUTIONAL QUANTIFICATION

By R. D. GALLIE

I

COHEN claims that Prior's main purpose in [1] was to argue that quantifiers binding sentence variables need not be objectual.

> To be thinking about a proposition, he [(Prior)] claimed (p. 29), is not to stand in a certain relation to an abstract object, and so to accept that there is some proposition about which Jones is thinking is not to accept the existence of an abstract object ([3] p. 69).

And indeed Prior does say on p. 29 of [1] that

> '— thinks that — is a foolish fellow' is a two-place predicate; its blanks keep places for *names of objects*.[1] But '— thinks that the proposition that — is an absurd one' . . . is not of this form; for here . . . the second blank *does not keep place for a name but for a sentence*.[1]

But Cohen immediately continues

> It is therefore rather misleading to apply Quine's substitutional/objectual dichotomy to the exegesis of Prior's theory if thereby Prior's theory has to be classified as objectual.

Now it is not at all obvious that if a quantifier that does not bind singular name variables fails the criterion laid down by Quine for being substitutional then it must be objectual in Quine's sense—a point that should have been stressed in [8]. Because to be objectual in that sense it has to bind singular name variables:

> Substitutional quantification in the substitution class of singular terms, or

---

[1] My italics.

> names, is the sort that comes closest to objectual quantification. But it is clearly not equivalent to it—not unless each of our objects is specifiable by some singular term or other, and no term of that substitution class fails to specify an object ([4], p. 105 f).

Perhaps Cohen thinks that all non-substitutional quantifiers bind singular name variables. How else could he accuse a critic who pointed out that Prior held that quantifiers binding sentence-variables were not substitutional of foisting upon Prior the view that such quantifiers would have to be objectual—i.e. would at least have to be quantifiers binding singular name variables?

Take the case of a logic whose particular quantifier binds common noun variables. (Such a logic might be an offshoot of syllogistic or the system ΣT2 of [2] chapter VII.) Of this quantifier Prior remarks ([2] p. 64)

> It is . . . misleading to translate, say, 'Σc' as 'There is an object c such that'. 'There is a kind-of-object c such that' would be more like it, but it is better to content ourselves with the simple 'For some c'. For even the translation 'There is a kind of object c such that' suggests that this kind-of-object is something which might *exist*, and this suggestion is misleading. These quantifiers . . . are like the English 'However' and 'Somehow . . .'; we do not expand 'Somehow' to 'There is a how such that', and neither should we expand the Σc of this system in any such fashion.

Now let the universe of discourse be denumerably infinite. Let F be a sentence-forming functor on common nouns and let us assume that the set of common nouns and common-noun-like constructions is also *denumerably infinite.* (Remember some common nouns will apply to some objects in the universe, some to all objects and some to no object.) It has been argued by Küng and Canty in [7] that in such a case

$$\ulcorner \Sigma c F c \urcorner$$

need not imply

$$\ulcorner F d \urcorner$$

for some common noun or common-noun-like construction 'd' because the extensions available for common nouns here constitute a *non-denumerable* set. That is, they have argued that the particular quantifier in question is not substitutional by Quine's criterion.

Plainly a common noun does not stand for an individual entity abstract or concrete over and above the entities or entity that it applies to, if any. Take 'horse' for example. 'Horse is a puzzling species' is syntactical nonsense (although 'The horse is a puzzling species' is not), and of 'Horses are a puzzling species' one is inclined to ask: what *one thing* does 'horses' stand for here? None of this is altered by the fact that in some universes there are 'more' extensions than there are common-nouns.

Again no doubt 'horse' has the class of horses as its extension. But in order to have it as its extension it does not have to denote it as a singular name denotes its denotation. And so the case for the non-substitutional nature of the quantifier in

$$\ulcorner \Sigma c F c \urcorner$$

does not include a case for changing the category of the variable bound by it.

Now let us turn to the case of sentence-variables. I do not know what Prior's reasons were for holding that the particular quantifier binding them was not substitutional. But clearly there is no case for holding in advance of knowing them that they must involve reasons for changing the category of the variables bound by the particular quantifier in question. In fact at this point it is worth noting that in certain cases the particular quantifier must be substitutional. For suppose all zero-place predicates have either the null class or the set of all n-tuples of objects in the universe for variable n as their extension in the case of classical predicate logic. Then a particular quantifier specially introduced to bind zero-place predicate variables could not fail to be substitutional since the class of extensions for sentences is manifestly outrun by the class of sentences. (Cohen's case for rejecting the Tarskian candidate for the extension of all true sentences rests on the desirability of replacing the usual truth-condition for conjunctive predicates by one in which the eliminability of 'true' on the right-hand side is, in general, lost.)

II

Granted that Quine says on p. 90 of [5] that

> The analogue of an unquantified "x" in ordinary language is a pronoun for which no grammatical antecedent is expressed or understood,

having previously said on p. 83 that

> the so-called existential quantifier '$(\exists x)$' corresponds to the words 'there is something x such that';

nevertheless he does not baulk at saying *this* on p. 85:

> Application of '(x)' to the expression:
>
> (10) x is identical with x
>
> in the fashion:
>
> (11) (x) (x is identical with x)
>
> is called universal quantification of (10) . . . . *To say that* (11) *is true is to say that no matter what object in the universe be imagined named by* 'x' *in* (10), (10) *becomes true.*[1]

---

[1] My italics.

So after all Quine does make use of specification or substitution or something very like that in introducing us to the meaning of *universally* quantified formulas. And presumably he would not think this passage incompatible with the one in which he introduced us to the meaning of existentially quantified formulas, particularly since he maintains (p. 86) that

> Universal and existential quantification are intimately connected in meaning, through negation

and that (p. 87)

> instead of writing '$(\exists x)Fx$' we could always write '$\sim(x)\sim Fx$'.

Moreover if unbound variables are like pronouns lacking antecedents then on replacing them by singular names one will get closed syntactically sound sentences. One cannot of course replace pronouns that do not lack antecedents and be sure of obtaining syntactically sound sentences. Surely this is all that Prior means by saying that unbound variables stand for, in the sense of keep a place for, names ([1] p. 35). He also says that a (singular nominal) variable stands for any object or person in the sense that it stands for (keeps a place for) any name that stands for an object or person ([1] p. 35). Once again one is hard put to see how even this is incompatible with the view that unbound (individual) variables are like pronouns without antecedents. They hold place for singular names which in turn stand for objects.

In other words apart from Quine's explicit insistence on an ontologically committed reading of the particular quantifier the differences between Quine and Prior when engaged in showing

> how the expressions and formulas of the system can be rendered coherently into idiomatic English

are grossly exaggerated by Cohen. Rather, with respect to his position on singular nominal variables, Quine is not really less 'substitutionalist' than Prior.

It may be said that the kind of reasons that Quine gives for rejecting his own substitutional characterization of quantifiers binding singular nominal variables are model-theoretic whereas Prior's are, by comparison, everyday. But what strikes one about the reasons is their similarity. Thus in [6] p. 92 Quine's argument is neatly summed up in the words

> In a generous universe *there are more things than can be named*[1] even with an infinitude of names.

Prior maintains that the truth of 'For some x, x is red-haired'

> *may be due to the red-hairedness of some object for which our language has no name*[1]

[1] My italics.

> or which no one is in a position to point to while saying '*This* is red-haired' ([1] p. 36).

And it is difficult to pinpoint anything in Prior's writings to stop him assenting to Quine's argument except his reluctance to view numbers as things (vide [1] p. 64).

What then is substitutionalism? Perhaps the key to its nature is to be found in the last paragraph of the first section of Cohen's paper:

> Prior's problem is to produce an interpretation for quantification binding propositional variables that is not of the 'there is something *p*' or 'there is a *p*' form. He . . . neatly side-steps the obstacles to solving his problem posed by standard interpretational procedures like those which Quine adopts in *Methods of Logic* . . . . The attempt to foist a non-substitutional interpretation of quantifiers onto Prior leaves it rather obscure how we are to understand Prior's neologisms 'anywhether' or 'somewhether', with the aid of which he proposes to translate quantification binding propositional variables.

Prior's neologisms are indeed an attempt to produce forms of words by which we may read such a quantifier as 'Σp' without ontological commitment to a realm of abstract entities. But such an enterprise, it has been argued in Part I, is only incompatible with accepting a non-substitutional semantics for a system for those who think that accepting such a semantics involves holding that its abstract objects have to be named by the parts of speech for which the variables of the system hold place. And so the fact that Prior did not discuss standard interpretational procedures in [1] does not mean that he saw their abstract objects as obstacles to his renderings of such quantifiers as 'Σp'. Rather, in certain cases the abstract objects acting as semantical counterparts to the variables of the system can even induce a substitutional characterization of the quantifier.

REFERENCES

[1] A. N. Prior, *Objects of Thought*, ed. P. T. Geach and A. J. P. Kenny, Oxford 1971.

[2] A. N. Prior, *Time and Modality*, Oxford 1957.

[3] L. J. Cohen, 'Roger Gallie and Substitutional Quantification', ANALYSIS 34.3, pp. 69–73.

[4] W. V. O. Quine, *Ontological Relativity and Other Essays*, New York, Columbia University Press, 1969.

[5] W. V. O. Quine, *Methods of Logic*, Second edition (revised), London, Routledge and Kegan Paul, 1962.

[6] W. V. O. Quine, *Philosophy of Logic*, Prentice Hall, Inc., 1970.

[7] G. Küng and J. T. Canty, 'Substitutional quantification and Lesniewskian quantifiers', *Theoria* XXXVI, 1970, pp. 165–82.

[8] R. D. Gallie, 'A. N. Prior and Substitutional Quantification', ANALYSIS 34.3, pp. 65–9.

University of Leicester

© R. D. GALLIE 1975

GODS AND VIRUSES

By Norvin Richards

IS religion just bad science, or is it something else altogether? It has been suggested that one function gods serve in religious thought is the same as that which theoretical entities serve in scientific theories: 'Like atoms, molecules and waves, then, the gods serve to introduce unity into diversity, simplicity into complexity, order into disorder, regularity into anomaly.'[1] I agree; but gods are also unlike standard theoretical entities in being conceived as unobserved *agents*. I shall argue that this accounts for the presence in religion of certain forms of life which are always foreign to the practice of science, and sometimes inimical to its purpose: worshipful attitudes and practices and disinterest in apparent disproof.

Consider first the virus, originally a theoretical entity invoked in order to explain apparent counter-instances to the theory that disease is caused by hostile organisms invading the body. Although the theory had many confirming instances, for certain diseases no bacterial agent could be isolated or identified. So it was postulated that in such cases the invading organism was a *virus*, something simply too small to be observed with existing devices. Since viruses were conceived as being otherwise very like standard invading organisms, ascribing puzzling diseases to them provided simplicity and order.

Understanding a puzzling event as the work of a god provides a similar simplicity and order. For, although unobservable, gods are always conceived as being like people in having perceptions, desires and intentions. Observable people are responsible for some events, just as observable bacteria are; so attributing other events to unobservable people simplifies in just the same way as attribution to unobservable bacteria.

Still, gods but not viruses are conceived as having perceptions, desires and intentions. What follows from that? First, it follows that gods, but not viruses, are suitable objects of the distinctly religious rituals of worship, prayer and sacrifice. For, it only makes sense to direct such rituals to objects you believe can observe and appreciate what you're doing. This point holds regardless of whether one takes the rituals to be informative to the gods or only as demonstrating something they already know. Of course, a community *could* perform formally identical ceremonies solely for other purposes: to bind its members, or as an aesthetic exercise or to earn money from tourists. But this wouldn't be worshipping, praying or sacrificing. The ceremonies would have lost their religious meaning and their title to those names, precisely because they would no longer be directed to a being believed to perceive and care.

---

[1] Robin Horton, 'African Traditional Thought and Western Science', *Rationality*, ed. Bryan Wilson (Oxford: Basil Blackwell, 1970), p. 134.

Worshipful attitudes and practices occur in religion and not in science because only in religion are the theoretical entities conceived as having perceptions, desires and intentions. And only such beings can be objects of worship.

It has often been suggested that the religious believer is unreasonably tenacious in his belief, irrational in being undisturbed by what is really very serious counter-evidence,[1] whereas the scientist is pictured by some as actively seeking falsification of his present theories in order to broaden them,[2] and by all as having a healthy respect for counter-evidence. One might expect such a difference: a god *mistakenly* discarded might well retaliate, after all, whereas a scientist can give up the concept of a virus or an atom without fear of retribution. Nor for the scientist changing theories the newly *lonely* universe of the convert to atheism.

One form the tenacity of religious belief takes is the relatively greater reliance on *ad hoc* devices to explain away counter-evidence. Horton (p. 149) offers the example of the Kalbari who posit three distinct sets of unobservable agents: ancestors, heroes and water-spirits. It is then easily explicable why an ancestor should act one way in one case and quite differently on similar occasions: inconsistency is simply blamed on interfering heroes or water-spirits. Similarly any mistaken predictions by the Azande poison-oracle are explained away as due to "impurity" in either the poison or those involved in the ritual.[3]

Of course, science is hardly devoid of *ad hoc* devices. Suppose a chemical succeeded in curing a viral disease on one occasion and failed on another. The inconsistency might well be attributed to an undetected impurity in the chemical or to an unnoticed fault in the application by the experimenter. But although this is *ad hoc*, at least its truth is testable in principle. This is not true of what the Azande mean when they call poison and practitioners impure. And while it is possible for a scientist to defend his theory in the Kalbari fashion, by invoking a hierarchy of interfering unobservables, if other things are equal, such a theory would be rejected in favour of one with a less complicated superstructure.

These differences are also due to the personal nature of gods and the impersonal nature of theoretical entities in science. Employing the concept of a god in one's explanatory framework requires divining the god's intentions, perceptions and desires. Only by assessing these can it be said whether a crop failure (say) was a punishment, a test or an oversight, and thus what response is needed to improve matters. Obviously the scientist needn't do this in order to employ the concepts of virus, atomic structure, wave, etc., just because these aren't conceived as agents.

---

[1] John Wisdom, 'Gods', *Proceedings of the Aristotelian Society, Supplementary Volume*, 1939; Antony Flew, 'Theology and Falsification', *University*, 1950–51; R. M. Hare, 'Theology and Falsification', *University*, 1950–51; Peter Winch, 'Understanding a Primitive Society', Wilson, *op. cit.*, pp. 78–111; Robin Horton, *op. cit.*

[2] Horton, *op. cit.*, reflecting Popper's well-known views.

[3] Winch, *op. cit.*

Now, we all know how difficult it sometimes is to know the intentions, perceptions and desires with which another human has acted. These are compounded infinitely when the agent is a god: one never sees the god performing the action, only its results. So we cannot learn what the god perceived by noting the position of his body, nor his intentions by observing his posture, his expression or how he moved while acting. All we have available are the reports of those who say the god has let them in on it and the broader patterns one thinks discernible in the god's other actions. These are obviously treacherous data. The reliance on predictive tools laced with *ad hoc* protections simply reflects the predicament one faces in interpreting the actions of an unobservable agent: nothing better is to be had.

This is all very well, perhaps, but where does it leave us with regard to the *large* question? Is religion bad science, or is it something else altogether? Suppose we define science as the effort to find regularities which will enable us to predict and control events. Religions are an effort to do this, although they also serve other purposes. Religion also contains forms of life—worshipful attitudes and practices and a special treatment of apparent counter-evidence—which are foreign to science and which can block the pursuit of manageable regularities in obvious ways. Explaining the presence of these forms of life in religion does not require invoking the other purposes religion serves or denying that it also serves the explanatory function of science or describing religious believers as perverse. I have argued instead that religions contain worshipful attitudes and practices and seem not to admit counter-evidence, because they employ as theoretical entities gods who perceive, desire and intend.

The University of Alabama © NORVIN RICHARDS 1975

UTTERER'S MEANING AND IMPLICATIONS ABOUT BELIEF

By MICHAEL CLARK

1. IT seems clear that sentences of the form '*p*, but I don't believe that *p*' are not self-contradictory. ('*p*' here functions as a schematic sentence-letter for sentences of the assertoric variety.) If Walker says, 'Inflation is getting worse, but I don't believe it', it might be true both that inflation is getting worse and that Walker does not believe this. Yet it seems equally clear that, taken literally, Walker's utterance is somehow self-defeating. For this reason, if Walker did actually make such an utterance, we should be inclined to find some other interpretation for it, for example that Walker was expressing his surprised recognition that inflation was deteriorating, that he was accepting something that he could scarcely bring himself to believe. But interpreted in this way the example loses its philosophical interest.

In saying that inflation will worsen Walker implies that he believes it will do so, but what he says does not logically entail that he has this belief. How could it, for his utterance contains no reference to its speaker, still less to his state of mind? And, if entailment were what is involved, we should be able to infer from my not believing what I say to its falsity; for entailments allow of contraposition.

Now all this is clear and very well-known. What is less clear is how it is to be explained. There is a vague feeling that saying (non-logically) implies believing because saying is essentially the expression of belief and no informative utterance could be possible unless our declarative utterances were typically sincere. But, whether or not this is true, one might reasonably ask for a more detailed explanation, and in this paper I want to show that Grice's account of utterer's meaning furnishes one. That will not in itself show that it is the right explanation, of course, since I shall not attempt to prove that Grice's account is correct: indeed, I shall suggest one respect in which even his latest available account is unsatisfactory. But the capacity of Grice's approach to explain the self-defeating nature of the utterances in question is nevertheless a point in its favour.

2. Grice's original account of utterer's meaning in 1957[1] yields quite a simple and plausible explanation of the self-defeating quality of '*p*, but I don't believe it'. For, on that account, for Walker to tell us that inflation is worsening is for Walker to make his utterance with the intention of inducing, by means of our recognition of his intention, the belief in deteriorating inflation. So, if he says, 'Inflation is worsening, but I don't believe it', he is intending to induce in us both the belief in

---

[1] *Philosophical Review*, Vol. 66, repr. in P. F. Strawson, ed., *Philosophical Logic*.

worsening inflation and the belief in his incredulity. But then how can he expect to induce the first belief merely by getting us to recognize his intention? For that would depend on our accepting his authority for the belief, on our accepting it because we knew he wanted us to and because we thought he wanted us to believe something he knew to be true. In trying to induce in us the belief in his incredulity, he is trying to undermine the very faith in his authority that we must have if he is to fulfil his first intention. Since he cannot but be aware that he is doing this, he cannot hope to succeed in fulfilling his first intention; and because he can't intend to do something he believes he has no chance whatever of succeeding in,[1] he can't really have both intentions. Walker cannot, therefore, make the utterance seriously and mean it literally.

We cannot accept this as a general explanation, however, since as Grice has shown more recently (1969, *op. cit.*), his original account of utterer's meaning is in need of elaboration. The examinee who writes ("utters") a sentence meaning it, does not usually intend to induce in his examiner a belief in the truth of what it expresses—he assumes his examiner already knows what he is being told. The man who breaks down and confesses that he has committed a crime may simply be admitting what he knows the police already believe. And so on. In the face of this sort of counter-example Grice has modified his account to make the 'direct intended effect' of a declarative utterance that *p*: that the audience should think that the utterer believes that *p*.

3. Grice's redefinition of utterer's meaning in 1969 is of some complexity, but happily we may ignore most of its clauses. The explanation of the eccentricity of '*p*, but I don't believe it' which we now get is itself only a little more involved and not without plausibility.

The 1969 article offers us a series of redefinitions purporting to lead up to one which is finally adequate. On the penultimate definition (IV), which it will be convenient to use first, the speaker who uttered a sentence of the form '*p*, but I don't believe that *p*' would have intended to induce in his audience the following beliefs:

(i) that he, the speaker, believed that *p*;

(ii) that he believed that he didn't believe that *p*.

Now it is perfectly possible for someone both to believe something and to believe he doesn't believe it. He may, for example, believe it unconsciously, or he may have deceived himself into thinking that he does not believe that *p* because he is ashamed of believing it. (These, indeed, seem to be the only main sorts of case.) In neither case will he be fully aware of his present cognitive state, so that anyone who confidently claims to be in such a state must be insincere. Since he can scarcely be unaware of the evident insincerity of his claim, he cannot hope to induce any audience to accept it, and therefore cannot have the relevant intention.

---

[1] As Grice himself says (*Philosophical Review* (1969), Vol. 78, p. 158): 'It is in general true that one cannot have intentions to achieve results which one sees no chance of achieving'. Cf. Wittgenstein's invitation to try meaning 'It's ten o'clock' by 'There are angels in heaven'.

For example, suppose a bishop has lost his faith but that he cannot face up to his atheism, since if he acknowledged it his life would seem to him a virtual sham. He therefore deceives himself into thinking that he is not really an atheist. He believes that God does not exist, but he believes that he does not believe this. He will scarcely want seriously to tell anyone else that he is an atheist who has deceived himself into thinking he is still a theist, since he will not admit to himself that this is his condition—if he did, his self-deception would have failed and it would not be true that he believed he was not an unbeliever. The episcopal self-deceiver is precisely a man who will not face up to his atheism, and, if he does, he ceases to be a self-deceiver. A man who tells us he is deceiving himself or has an unconscious belief that *p* cannot expect to be accepted as sincere and so cannot intend us to believe him.

4. The foregoing explains the oddity of utterances of the form in question, as well as some of the other utterances that exercised G. E. Moore,[1] when they are made to some audience. But it is equally odd to think or say such things to oneself, or, for example, to write them in a private diary. This constitutes a general difficulty for 'communication-intention' theories of meaning, of course, and is not specifically generated by the examples under discussion. Grice's detailed and final revision of his definition, calculated to embrace utterances in the absence of an audience, will be found on pp. 175–6 of his 1969 article. Once again I shall avoid the full detail of his account and simply deal with the critical amendment. The utterer now intends his utterance *to be such that* anyone who ϕ would think that he the utterer has the belief involved. We are told that appropriate substituends for 'ϕ' are such expressions as 'is a passer-by', 'is a native English speaker', 'is identical with Jones'. Silent thinking is to be treated as if it were speaking aloud. ('We could perfectly well . . . replace . . . every process of speaking to oneself by speaking aloud or writing.' Wittgenstein, *Blue Book*, p. 4.)

But there seems to be a fairly obvious difficulty with this amendment. I might surely say something to myself or write it in my private diary for my personal satisfaction, perhaps whispering it in a distorted voice, or writing it in a private code, intending it to be such that *no one* who might overhear or see it would think I believed it. A man might say or write, 'I wish my wife would drop dead', intending his utterance to be such that no possible audience would realize he thought that, should he be unlucky enough to have any sort of audience for what was meant as a wholly private remark uttered to vent his feelings. (If this involves a misunderstanding of what Grice intends by the utterance's '*being such that* an audience would think . . .', I do not see what he can mean.)

---

[1] See his *Philosophical Papers*, 'Certainty', pp. 227–8.

Now, if such an objection were fatal to communication-intention accounts of meaning, one could not use them to explain the case of '*p*, but I don't believe it'. Fortunately, however, the objection is not at all fatal. True I cannot think of any way of amending Grice's final definition, but I shall argue that it is not necessary to do so in any case. The leading idea behind Gricean accounts of meaning is that language is essentially a public vehicle of communication and could not be developed except in a way that made it suitable for such communication. The inspiration has an obvious Wittgensteinian source, however unfaithful to the spirit of the later Wittgenstein the resulting theory, with its precise necessary and sufficient conditions. Once one has learnt a public language, one can then use it for private, non-communicative purposes. (Perhaps it could have been used like that all along, provided it had been developed in such a way that it was in principle teachable, but in practice this is unlikely.) The non-communicative uses will be parasitic on the communicative, in the sense that the former will be possible only because of the existence of the latter. It should be no surprise that no definition is easily found to cover both the standard communicative and the parasitic non-communicative uses, in the manner Grice is seeking. Surely it should be enough if the standard cases can be captured by a definition. In much the same way the concept of an unconscious fear is parasitic on that of a conscious one. In explaining what it is consciously to fear something it must be made clear that the subject is aware in some way of his emotion and of its object; but this essential feature will be lacking in the case of an unconscious fear. It will therefore be impossible to give a unitary, non-disjunctive definitional account of fear that embraces both conscious and unconscious cases; but to expect this is in any case to ask too much. Likewise, if Grice can give an account of utterer's meaning for the standard cases where the utterer has some audience, it should not be regarded as a deficiency of his account that his definition cannot readily be modified to embrace the parasitic cases. And, in the specific cases that are the concern of this article, an explanation of the deviance of utterances of the form '*p*, but I don't believe that *p*' when they are addressed to some audience is reasonably all we should ask for. True, some utterances are deviant in the absence of an audience too, but, if the Gricean approach to meaning is tenable, such utterances will only be possible in the absence of an audience if they are possible in the presence of one. If we can explain why they cannot satisfactorily be made in the presence of an audience, we need no further explanation of their unsatisfactoriness in the absence of one.

University of Nottingham

© MICHAEL CLARK 1975

AN OBSERVATION ON ENGLISH TRUTH-FUNCTIONS

By KEITH HALBASCH

HOW many truth-functional words (or morphemes) are there in English? In this paper I will prove, by a simple logical argument, that there are more than are usually suspected. I will also briefly point out and comment on some problems that this discovery raises for empirical linguistics and make some suggestions, though I shall attempt no detailed solution.

It is usually suggested that the truth-functions for which there are truth-functional words in English are some proper subset of the twenty possible unary and binary truth-functions. However, English allows constructions of the form ⌜Either ϕ_1 or ϕ_2 or ... or ϕ_n⌝, for any n. But the evidence here for the existence of a truth-functional expression of polyadicity greater than 2 is not conclusive, since such sentences can be construed as a series of binary disjunctions. Since the truth-conditions would be the same either way, semantic evidence cannot be employed to decide the case. However, the situation is not so bleak for a closely related case, viz., ⌜Neither ϕ_1 nor ϕ_2 nor ... nor ϕ_n⌝. Sentences of this form, for any n, are intuitively true just in case each constituent is false. Consider now the case where $n=3$. One might think that such sentences could be viewed as containing two binary 'neither ... nor''s, with one of the 'neither''s having been deleted. But this time the semantic evidence is telling; the truth-conditions are wrong! There are two ways of bracketing such a sentence (⌜(Neither ϕ_1 nor ϕ_2) nor ϕ_3⌝ and ⌜Neither ϕ_1 nor (ϕ_2 nor ϕ_3)⌝), but neither is truth-functionally equivalent to the corresponding sentence consisting of the same component sentences and the ternary 'neither ... nor ... nor'. Reconstruing the order of the sentential components will not help either, since ⌜(Neither ϕ_3 nor ϕ_1) nor ϕ_2⌝ is not truth-functionally equivalent to the sentence we are after either. Other reorderings are also of no use since the binary 'neither ... nor' is commutative. A quick proof that all combinations fail is that they are all false for the case where all three component sentences are false, whereas the ternary 'neither ... nor ... nor' yields a true sentence in this case. Proof: whichever two components are combined first, the combination is true; but this combined with the remaining false component yields a false sentence. Of course, since binary joint denial is a truth-functionally complete mode of sentence composition it will be possible to find a sentence containing just binary 'neither ... nor''s and occurrences of the original sentence types which will be truth-functionally equivalent to the sentence based on the ternary 'neither ... nor ... nor', but these are unintuitively complicated and all involve more than one occurrence of, at least, some of the basic components. I think that syntactic grounds alone rule these out of consideration.

The case is similar for $n>3$, but more difficult to prove generally because of the rapidly growing number of ways in which the components might be bracketed. For $n=4$, none of the 14 semantically distinct ways of combining the components yields a sentence equivalent to the corresponding one based on a 4-ary 'neither . . . nor . . . nor . . . nor'. The quick proof does not go through here since some of the combinations do get the all false case true as it ought to be. For example, nesting to the left will do for this case, but it will get other cases wrong. (E.g., the true-true-true-false case.) I have discovered no quick general proof for $n=4$ and have had to rely on the laborious use of truth-tables. For this reason I have not proceeded to any cases where $n>4$. Nevertheless, it is easy to see that some of the most intuitive methods of combination will not work in general. For example, nesting to the left will not work for odd n's since in the all false case it always produces a false sentence, where, of course, intuitively a true one is required. This holds regardless of the order the original components are taken in. Since some kind of nesting is extremely plausible from a syntactic point of view and since no ordering will work for $n=3$ and $n=4$, I conclude that it is highly probable that, for every positive integer n, English contains an n-ary joint denial operator. By analogy, we may want to arrive at similar conclusions for disjunction and perhaps conjunction and even other truth-functional connectives.

These conclusions entail that the truth-functional structure of English is more complicated than normally supposed. The truth-functional vocabulary will have to be specified recursively to start with or be in part derived by transformation at a later stage. The latter possibility seems most plausible to me. For example, we could employ a transformation that took us from sentences of the form ⌜It is not the case that (ϕ_1 or ϕ_2 or . . . or ϕ_n)⌝, where each 'or' is construed as binary, to ⌜Neither ϕ_1 nor ϕ_2 nor . . . nor ϕ_n⌝ where we have a single n-ary connective. The generality achieved here by use of ellipses is important. A more rigorous procedure would perhaps specify infinitely many precise transformations by a recursive procedure. The simple fact is that the results arrived at above require a significant complication in our treatment of truth-functions in English.

University of Connecticut

© KEITH HALBASCH 1975

ANALOGY, INDUCTION AND OTHER MINDS

By THEODORE W. BUDLONG

CAN the argument from analogy for the existence of other minds be formulated in such a way that the 'generalizing from a single case' objection is avoided? A way of trying to do this, proposed by Ayer, has been gaining acceptance. I think it is faulty, and want to undermine it before it becomes further entrenched.

A person using the argument from analogy argues from associations he has discovered between his own behaviour and mental states to an association between behaviour and mental states on the part of others. To take a common example, I find that my pain behaviour is normally accompanied by pain. I then infer, by analogy, that similar behaviour on the part of others is accompanied by pain.

The move that Ayer suggests is this (*The Problem of Knowledge*, Baltimore: 1956, p. 219):

> The objection that one is generalizing from a single instance can perhaps be countered by maintaining that it is not a matter of extending to all other persons a conclusion which has been found to hold for only one, but rather of proceeding from the fact that certain properties have been found to be conjoined in various contexts to the conclusion that they remain conjoined in further contexts.

Following Ayer's lead, Alvin Plantinga and Michael Slote have formulated the argument as an enumerative induction in the following way (*God and Other Minds*, Ithaca: 1967, p. 247; *Reason and Skepticism*, London: 1970, p. 113):

> (A) Every case of pain behaviour such that I have determined by observation whether or not it was accompanied by pain in the body displaying the behaviour in question *was* accompanied by pain in that body.
> Therefore, probably every case of pain behaviour is accompanied by pain in the body displaying it.

Plantinga and Slote find nothing *prima facie* wrong with gathering all the evidence for the premise of (A) from one's own case. Although they disagree about the argument's strength, they both believe that if the argument fails, it fails for other, more rarified, reasons.

It is here that they go astray. In an enumerative induction, one argues from having found a number of *P*'s to be *Q*'s to the conclusion that probably all *P*'s are *Q*'s. Such an argument will be a weak one if the observations are confined to a relatively small number of *P*'s or, even worse, to only one *P*. It is a fallacy to suppose that repeatedly observing a single *P* is a satisfactory substitute for examining a number of different *P*'s. Such repeated observations may support a generalization about that one individual, but not a generalization about all *P*'s. This is the mistake made by Plantinga and Slote. They want to generalize about correlations between pain and pain behaviour on the part of all persons on the basis of repeated observations of one person.

The defence of (A) suggested by Ayer's remarks is that the argument is not about persons but about instances of pain behaviour. That changing the subject in this way is specious becomes apparent when we compare it with a similar move that could be made by Maxwell, a lazy grocer. Faced with the charge that his milk is sour, Maxwell opens only one of the quart cartons in his store. Drinking it all, and finding it to be in good condition, he declares that all the other cartons are also fit for sale. When it is asserted that Maxwell should check some other cartons, he replies that he has performed a perfectly satisfactory enumerative induction. He has applied his induction, he says, not to cartons of milk but to ounces of milk. Having drunk an entire quart, he has examined thirty-two fluid ounces of milk, and concluded on the basis of observing thirty-two instances that every fluid ounce in his dairy case is in good condition.

Maxwell and the defender of (A) claim to have examined a large number of instances, but their observations are redundant. Consider the situation of Maxwell after he has consumed two ounces of milk from the carton he has opened. His drinking a third and a fourth ounce will not strengthen his claim that every ounce of milk in his store is in good condition; drinking from that carton after the first sip can only satisfy his thirst. Similarly, observing still another instance of one's own pain behaviour, and finding it to be accompanied by pain, will not make it any more probable that the pain behaviour of other persons is accompanied by pain. At best it can only reinforce the claim that this is true in one's own case, a claim that for most of us hardly needs reinforcing.*

---

* I thank Robert Jaeger and William Rowe for helpful comments and criticism.

Purdue University

© THEODORE W. BUDLONG 1975

26 [illegible] 1975

NOTES

The Analysis Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Margaret A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, A. C. MacIntyre, Bernard Mayo, D. H. Mellor, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

Subscriptions. The subscription to Analysis is £2.50 ($7.50) for individuals in Great Britain, £3.00 ($9.00) for individuals abroad, and £3.50 ($10.50) for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford, or placed with any bookseller.

Contributions. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgment is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in Analysis remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

26 DEC 1975

PRINTED IN GREAT BRITAIN BY BURGESS AND SON (ABINGDON) LTD

l. 35 No. 4 (New Series No. 166) March 19

ANALYSIS

Edited by
C. J. F. Williams

UNIVERSITY LIBRARY

CONTENTS

Forgiveness — Martin Hughes

Classical semantics and entailment — Larry Briskman

Adjunction and paradoxical derivations — Charles F. Kielkopf

Logical subtraction — James L. Hudson

Confusing the audience — G. H. Bird

Are "Cambridge" changes non-events? — Paul Helm

BASIL BLACKWELL . ALFRED STREET . OXFORD

Sixty pence

ANALYSIS 35.4 MARCH 1975

FORGIVENESS

By MARTIN HUGHES

IN reply to Mr. Gingell ('Forgiveness and Power', ANALYSIS 34.6, pp. 180–3), I argue that Christians who speak of God as all-forgiving show an excellent understanding of what forgiveness is.

Forgiveness is the cancellation of deserved hostility and the substitution of friendlier attitudes. It has important consequences, for which it is highly valued—socially, in that the offender can hold up his head again, and inwardly, in the quietening of remorse.

Forgiveness has a discretionary nature, and the discretion belongs to the injured. This aspect of forgiveness may make it seem arbitrary and unreasonable, and there may be doubt whether the important and valuable consequences mentioned really should flow from it: why should my life be easy because I have injured an indulgent and soft-hearted man, and yours miserable because you have injured (perhaps in the same way) one who is stern or spiteful?

Rational moralists have long found this a severe problem. Seneca, who wrote an essay *On Clemency* addressed to Nero of all people, was very well aware of the discretionary nature of forgiveness in the entourage of an absolute monarch, where even formal justice was a personal matter and might be highly tinctured by transient emotions (Loeb edition, pp. 414, 430)—yet he makes an interesting effort to combine discretion with rules, only to end up with a complex contradiction: it is a rule that we always use discretion significantly to reduce penalties (both public and personal, as I take it) from what, under the rules, is due (ibid., pp. 434, 448). Butler, in his ninth Sermon, seems to bury the discretion of the injured under an alleged duty to assimilate one's feelings to those of a disinterested but virtuous bystander. On this sort of view, the principle that only the injured can forgive becomes, at best, a linguistic triviality: all that is morally important cleaves to the ideal observer. Butler might then have a kind of answer to Mr. Gingell, but it would leave him with a very etiolated Christianity, in which the sense of personal involvement of God with the sinner was lost. But Butler's view is quite coherent. If we treat the discretion of the injured as morally important, we express a moral conviction, not a logical insight.

Discretionary forgiveness cannot fully be regulated by definite and universalized moral rules. If I adopt such rules, I cannot treat others as having the right to forgive or not at their discretion when they are injured, but only as having the duty to forgive or not according to my rules: in a way, I make myself the ideal observer. But completely unfettered individual discretion may become a matter of temperamental caprice, as Seneca feared. Yet discretion remains discretion and does not become caprice if it is receptive of informal advice. If someone advises me formally, he gives me orders: he propounds deductions from rules he thinks I accept, or is certain I should accept, and I am under threat of condemnation if I do not comply. If someone advises me informally, he understands and condones my doubts about certain rules, or sees that we may not be able to show rigorously how accepted rules apply concretely. Still, we often seek informal advice "to clarify the issue": we see at least what the rules are we have to consider, and this forces us—since otherwise it is very hard to see the real import of abstractly formulated rules—to think of similar cases and the decisions of others: thus our discretion loses its pure individualism and its tincture of caprice, since it comes to take serious account of the experience, maybe the wisdom, of others. We usually find that there are others whose decisions we respect—even if, or especially if, we do not believe we would have chosen so well in their place. If we can, it is to people we so respect that we go for informal advice, even though we know they cannot give us orders on matters for our discretion: conversation with them will bring home to us with unlessened force the fact that the decision is ours.

Someone can achieve general respect for the advice he gives, and there may be contact between him and the wrongdoer as well as the injured. This is illustrated by Mr. Gingell's case of the penitent son, who has brought injury on others while disobeying his father's instructions. He doubts whether some of the injured will forgive. His father says 'The important thing is that I forgive you', thus appearing a monster of egotism. But he may sound less monstrous if he says 'You have done all you could be asked so as to make amends: now you should not torment yourself so much'. This may be the right thing to say, and if he has acquired and deserved his son's respect it may be important that it is he who says it. Such informal advice is most often valued, not because it can be supported by demonstrative arguments or appeal to uncontroversial first principles—which usually it cannot—but because of our respect for the person who offers it. Coming from someone known to be over-indulgent it would be little regarded.

If the father helps conquer his son's remorse, his words share some of the effects of forgiveness. But is there reason to say that the son is now *forgiven,* rather than just relieved? Yes, since in some circumstances the discretion to forgive granted by the moral system may, under the same system, lapse. Thus the unreasonably embittered man finds in the end that no one cares what he thinks, and that, even if he now wishes, he cannot take part in the sort of act of forgiveness that was at first open to him: by now it is he who is in the wrong. The best available representatives of the moral system, like the father in the story or the social group in common fact, take over the right to rehabilitate the wrongdoer. We may now permit the injured to be bypassed, since his is not now *deserved* hostility: it is this feature of our moral system which is indicated by the use of the word 'forgive' for certain acts of external authority. And this feature of our moral system is a natural consequence of the fact that discretion can be genuinely granted, yet expected—as is normal with political discretions—to be used with some regard for certain ideals.

The right of individuals to represent the ideal is not conceded too often. This is because the authority of any social agent must be very questionable. If he is disinterested, he may be too detached and unsympathetic; if he is involved, he may find his judgment unbalanced. And he may be misinformed. So clear superiority of wisdom over those involved is hard to establish, unless they show some especially deep degree of folly or unreason. Only with God is there an answer to all these doubts. God cannot be misinformed, and more important, he personifies *both* personal involvement in love *and* the most high and disinterested justice. Jesus, my fellow-man and friend, stands for the former ideal. The latter is represented in the Old Testament by the God whose face man cannot see and live, and who surrounds himself always—most memorably when giving his Law—with fires and storms, the symbols of punishment and destruction; in the New Testament by the God who is so far untouchable and unswayable by personal feeling for us that he must punish the world's disobedience to the Law by the world's total destruction, except insofar as mercy is gained by a Mediator who offers himself for destruction in man's place.

To say that Jesus is God is to unite these two ideals. The union seems to contain irretrievable contradiction, but is just what we need if we are to guard God's authority against *all* criticism. If God's forgiveness seems too much like indulgence, we are to think of his transcendent justice; if God in his transcendence seems too forbidding and remote, then we are to think of his universal love. Only so can the penitent and the outcast be assured of hope when others, including those he has injured, spurn him. For God, as the *ideal* moral agent, always has the right to declare that they have abused their discretion, and that he forgives in their default.

If Christians use the same verbal form for human and divine transactions, when the rules governing usage cannot be exactly the same in both cases, this to me is a fact to be explained. To Mr. Gingell, it means that one usage enshrines a mistake. But such 'double usage' is not the same as contradiction: it simply tells us that certain connections are habitually made. The connections here are between the ideas of real and of ideal moral agency, and they reflect the fact that Christianity understands what kind of ideal authority can bring peace of mind to the wrongdoer—not badly, but all too well. The fault is in the Christian idea of God, not of forgiveness.

The double usage suggests that man is like God, enough for imitation. This suggests in turn that it is worth a man's while to hold up as an ideal to himself an image of all the characteristics he would need in order to take certain decisions well. Two points arise here from the Euthyphro problem, to which Mr. Gingell aptly refers. *First,* the acceptance of such an ideal image is not necessarily either an act of self-enslavement or a furtive attempt at self-worship through the fiction of a Master's Voice which is really one's own echo. If someone has a loyalty or commitment to an authority—that of a group or that of a single person—we do not necessarily have any reason against accepting this as an expression of autonomous morality: indeed any kind of moral life lacking all such commitment seems unimaginable. On the other hand, when someone refers to the characteristics needed to take decisions well, he need not mean 'to take them as I would take them' or even 'in some way I could specify': otherwise, the common experience of recognizing the limitations of one's own judgment without being certain how to improve it, and of recognizing another's judgment as superior, could not occur. This commendation by respect is logically distinct from immediate commendation: one may follow the other judgment even though well aware of a personal and immediate judgment of the situation which is quite contrary. This is commitment not enslavement, because it is clearly a form, not a negation, of personal and autonomous choice; it is self-discipline not self-worship, because the other judgment plays a non-redundant part. Accordingly, both propositions offered by Socrates to Euthyphro may be asserted: the Gods command what is good (as good is recognized by my autonomous judgment) because it is good, and yet I may reasonably think something good because they command it. In other words, I commend the Gods by respect and I sometimes prefer the judgment arising from this commendation to that arising from immediate commendations or opinions in certain situations. A moral exemplar must be comparable to me, only much better, and so my God is in my image, an idealized self the thought of whom sometimes has influence on my real unideal self. So. when an offence is committed, it outrages the morality I should, ideally, cherish: if I am a religious man, I must say it outrages my God, and I am only being consistent if I add that the forgiveness of God is what, above all, the wrongdoer needs. For, to me, all that is of most moral importance in matters of forgiveness is bound up with the readiness of all conerned to use their discretion, maybe contrary to their immediate opinions, in accordance with a certain ideal.

Hence the *second* point, which is that, when it comes to moral assessment of these religious ideals, the fictions and contradictions which often inhere in them need not much distress us. The commendation of an ideal of forgiveness is a form of advice in a discretionary matter, and such advice must, by my previous argument, be informal. Here myths come into their own. Myths usually contain fiction and contradiction, but still serve their purpose, which is to give to a certain kind of understanding—here moral—such articulation as it will bear, and still have importance, which is in fostering certain kinds of action. Euthyphro was bad at rational argument, but could still use his anthropomorphic fictions to arrive, for good or ill, at an unconventional moral certainty. The Christian view of the all-forgiving God is on my account contradictory, but has promoted, for good not ill, mercy and forgiveness among men—at any rate on the level of informal personal morality where religion belongs.

Plato thought that myths were only for minds in tutelage, and could always in the end be replaced by rational understanding of the good. My argument has been that in matters of discretion strictly rational morality fails; if rational theology is drawn in, it fails likewise. The consequent recourse to myths, even if they are myths of a divine lawgiver who seems, happily to the penitent or ineptly to the philosopher, to pre-empt our proper discretion, is really an acknowledgment of these failures and hence of our discretion. But it is also the best way of reminding us of certain ideals, which illustrate excellence in the use of discretion, and hence intimate what abuse would be. So they may also remind us that discretion, if abused, can be discredited or lost.

University of Newcastle upon Tyne

© Martin Hughes 1975

CLASSICAL SEMANTICS AND ENTAILMENT

By LARRY BRISKMAN

IN a recent paper, John Cleave [2] offers a non-classical account of entailment which is, nevertheless, based upon the classical notion of logical consequence (i.e., that B is a logical consequence of A iff B is true in all models of A, or $A \supset B$ is a tautology). Rather than placing restrictive conditions on the semantics of entailment, as done, for example, by Geach [4] and von Wright [7], Cleave places restrictions on the type of formula which can enter into an entailment relation. As I shall show, Cleave's notion of entailment does not adequately capture the intentions of entailment theorists. But this is not to say, of course, that the Geach–von Wright theory is free from blemish. In fact, I shall raise an objection to their theory as well. The upshot of the discussion will, in essence, be a defence of the classical account of entailment.

It is important to remember at the outset that non-classical entailment theorists have not been concerned *primarily* with avoiding the so-called paradoxes of strict implication. Rather, they have taken these "paradoxes" as refuting the claim that the intuitive notion of logical consequence, or of B's following logically from A, can be identified with either strict implication or the (equivalent) classical, Bolzano–Tarski notion of logical consequence.[1] In other words, the primary concern of entailment theorists has been to capture (formally) the intuitive notion of logical consequence which, it is claimed, is *stronger* than the classical notion, and for which the classical notion provides *necessary*, but *not sufficient*, conditions. It is, in fact, quite easy to see how this concern can arise. If we start from the assumption, unquestioned by almost all writers on entailment, that mathematical truths are necessarily true, then an identification of entailment with strict implication leads to the seemingly absurd consequence that all mathematical truths entail each other, and so are logically equivalent, and thus, assuming our derivation rules are complete, inter-deducible. But this makes a nonsense of mathematical (and logical) practice; for in mathematics (and logic) we usually prove theorems from axioms and can only rarely reverse the procedure and prove the axioms from theorems. Moreover, assuming that mutual entailment is a criterion of propositional identity, it would follow that there is, at base, only *one* mathematical truth, which seems absurd. In addition, consider the following two mathematical truths: (1) All even numbers greater than 2 and less than or equal to 100 are the sum of two prime numbers, and (2) All even numbers greater than 2 and less than or equal to 200 are the sum of two prime numbers. It certainly looks as if, although (2) entails (1), (1) does *not* entail (2). But if (1) and (2) are both necessarily true, and entailment is identified with strict implication, then (1) *does* entail (2). Hence, it is claimed, the notion of valid proof and logical consequence used in mathematics (and logic) cannot be identified with strict implication. Non-classical entailment theories are meant to make good this defect; and a rational evaluation of any such proposed theory will thus crucially turn upon how well it solves this problem.

---

[1] I should perhaps mention here that Lewis's strict implication (p strictly implies q iff not possibly (p & $\bar{q}$)) is only equivalent to the Bolzano-Tarski notion of logical consequence if we restrict the modal operator to ranging over allowable interpretations or models. For in that case 'not possibly (p & $\bar{q}$)' simply reduces to 'there is no interpretation in which (p & $\bar{q}$)' and therefore no interpretation which models p but fails to model q. This shows, by the way, that modal notions are completely unnecessary in the characterization of 'logical consequence'. In this paper I shall use 'strict implication' as an equivalent expression to 'classical logical consequence'. Incidentally, it might be retorted that since 'allowable interpretation' just means 'possible interpretation' modal notions have not been avoided. This is not so: 'No A's are B's' certainly looks like a *possible* interpretation of 'All A's are B's', but it is not an *allowable* interpretation. Classical logical consequence is relative to allowable interpretations, not possible ones.

In the light of this, we now turn to Cleave's account of entailment. Cleave's theory makes use of the notion of a rigid formula: a formula F is rigid iff it does not contain an *irrelevant* subformula X, and 'an occurrence of a subformula X of the formula F is *irrelevant* if the logical equivalence class of F remains unchanged when that occurrence of X and only that occurrence is replaced by its negation' ([2], p. 119). In other words, if $F(X) \equiv F(\bar{X})$ then the occurrence of X in F is irrelevant, and so F is not rigid (i.e., F is slack); while if $F(X) \not\equiv F(\bar{X})$ for *all* subformulas X of F then F is rigid. After investigating some of the properties of rigid formulas, Cleave suggests the following definition of entailment: $A \rightarrow B$ (or *A entails B*) if, and only if, $A \supset B$ is a rigid tautology. It is clear that the rigid tautologies form a subset of the class of tautologies, and hence that strict implication becomes a necessary, although not sufficient, condition of an entailment. Thus, Cleavian entailment is a stronger notion than is strict implication, even though from a semantic point of view it makes use of the same notion of consequence.

The intuitive idea behind Cleave's suggested definition is not hard to find. If $A \supset B$ is a rigid tautology, then the antecedent contains no subformula which is gratuitous for the "deduction" of the consequent. That is, no component of the antecedent can be negated without destroying the tautologousness of the conditional. Equally, if $A \supset B$ is a rigid tautology, then the consequent contains no subformula which has been gratuitously "deduced" from the antecedent. That is, no component of the consequent can be negated without again destroying the tautologousness of the conditional. Thus we can say that rigid tautologies correspond to logically "tight" arguments, with no superfluous "noise" (i.e., informational redundancies). As Cleave rightly points out, 'most, if not all, of the commonly agreed paradoxes of implication arise from slack [i.e., non-rigid] tautologies' ([2], p. 118). What Cleave fails to point out, however, is that many, if not most, commonly agreed valid patterns of argument and proof fail to correspond to rigid tautologies, and hence to Cleavian entailments.

First, consider the following three tautologies: (1) $(p \supset \bar{p}) \supset \bar{p}$; (2) $(p \supset (\bar{r} \,\&\, r)) \supset \bar{p}$; (3) $(p \supset (\bar{p} \,\&\, p)) \supset \bar{p}$. Now, while (1) and (2) are rigid, (3) is not rigid since $(p \supset (\bar{p} \,\&\, \bar{p})) \supset \bar{p}$ is also a tautology. Hence, $(p \supset \bar{p}) \rightarrow \bar{p}$, $(p \supset (\bar{r} \,\&\, r)) \rightarrow \bar{p}$, but *not* $(p \supset (\bar{p} \,\&\, p)) \rightarrow \bar{p}$. These examples show, first of all, that Cleavian entailment is not closed under substitution: for while (3) is a simple substitution instance of (2), (2) answers to an entailment, but (3) does not. This already makes the identification of entailment with the rigid tautologies highly suspect, since failure of closure under substitution limits severely the possibility of *unrestrictedly* valid "principles of inference" corresponding to valid entailments. Thus, take the so-called Principle of Disjunctive Syllogism: from $A \vee B$ and $\bar{A}$, infer B. Now while (4) $((p \vee q) \,\&\, \bar{p}) \supset q$ is a rigid tautology, (5) $((p \vee \bar{p}) \,\&\, \bar{p}) \supset \bar{p}$ is not rigid since $((\bar{p} \vee \bar{p}) \,\&\, \bar{p}) \supset \bar{p}$ is also a tautology. In other words, Disjunctive Syllogism, and hence the Law of Detachment for material implication, is not unrestrictedly valid in Cleave's system precisely because Cleavian entailment is not closed under substitution.[1] Secondly, (2) and (3) land us in the odd situation that p's materially implying any contradiction *other* than $(\bar{p} \,\&\, p)$ entails $\bar{p}$, but p's materially implying $(\bar{p} \,\&\, p)$ *does not* entail $\bar{p}$. This seems arbitrary at best. Moreover, (1) shows us that had we *stopped* the argument at p's materially implying $\bar{p}$ (rather than $(\bar{p} \,\&\, p)$) then we *could* have validly inferred $\bar{p}$. This seems to make the validity of an actual argument hinge upon a kind of historical "accident". In addition, arguments corresponding to (3) seem (intuitively) to be perfectly valid arguments. Yet they do not count as valid arguments in Cleave's system.

Even more damaging, however, is that Cleavian entailments hold only between logically contingent propositions (see Theorem 5, p. 121). Thus, *reductio ad absurdum* proof becomes strictly speaking impossible since no premiss (i.e., antecedent) *can* entail a contradiction. Moreover, even though Cleave's entailment is restricted to logically contingent propositions, it is still nevertheless *not* unrestrictedly transitive (see his example 3, p. 119). This is "paradoxical" for the following reason: the so-called paradoxes of strict implication arise in precisely those cases where the truth of 'not possibly (p and $\bar{q}$)', and hence of 'p strictly implies q', is determined by the impossibility of one of the conjuncts *alone*. Seemingly, if we restrict our attention to logically contingent propositions, then the truth of 'not possibly (p and $\bar{q}$)' *cannot* be so determined, and therefore entailment reduces to strict implication,[1] which *is* unrestrictedly transitive. Thus the "paradoxical" result is obtained that Cleave's entailment *is* restricted to logically contingent propositions and yet *is not* unrestrictedly transitive and so *does not* reduce to strict implication. What this argument shows is just how restrictive Cleave's entailment notion is, for it is more restrictive than strict implication even in that area where the latter is *immune* to "paradox". It would seem, then, that Cleave's theory pays much too high a price for the avoidance of the so-called paradoxes which "beset" strict implication.

---

[1] In fact, there can be *no* unrestrictedly valid "principles of inference" in Cleave's system for the simple reason that Cleave's entailment notion holds only between logically contingent propositions. Thus substituting either $(p \vee \bar{p})$ or $(p \,\&\, \bar{p})$ for the metavariables A, B in '$A \rightarrow B$' will destroy the entailment. But my examples (2) and (3), and (4) and (5), show that closure under substitution fails even if we *restrict* our substitutions to logically contingent propositions.

To see just how high is the price, and thus how inadequate the rigid tautologies are as an explication of entailment, consider the famous Lewis argument. This argument is designed to show that *any* theory of entailment is bound to be "paradoxical", in the sense that any theory which denies the "paradoxical" entailment $(p \,\&\, \bar{p}) \rightarrow q$ can only do so at the cost of denying some intuitively *desirable* property for entailment. Thus the following seem to correspond to "valid modes of inference" or valid entailments:

(1) $A \,\&\, B$ entails A
(2) $A \,\&\, B$ entails B
(3) A entails $A \vee B$
(4) $((A \vee B)$ and $\bar{A})$ entails B
(5) $((A$ entails $B)$ and $(B$ entails $C))$ entails $(A$ entails $C)$.

But on the basis of these posited properties of entailment we can prove the "paradoxical" $(p \,\&\, \bar{p}) \rightarrow q$ as follows:

(a)	$p \,\&\, \bar{p}$	premiss
(b)	p	from (a), by 1
(c)	$\bar{p}$	from (a), by 2
(d)	$p \vee q$	from (b), by 3
(e)	q	from (c) & (d), by 4
(f)	$(p \,\&\, \bar{p}) \rightarrow q$	from (a)–(e), by 5

Now most entailment theories attempt to break this proof-chain at, at most, one point. The reasons for this are clear. First, one wants to preserve as many entailments as possible while avoiding the "paradoxical" (f), and so does not want to deny more of (1)–(5) than is absolutely necessary. Secondly, one does not want a notion of entailment which is even more "paradoxical" than is strict implication itself. Thus, for example, Anderson and Belnap [1] reject principle (4); Geach [4, 5] rejects transitivity, or (5); while the four systems investigated by Smiley (in [6]) deny successively either (1 & 2), (3), (4) or (5). What is intriguing is that in Cleave's system *all* of (1)–(5) fail: (1)–(3) fail unrestrictedly, while (4) and (5) do not hold unrestrictedly, even if restricted to logically contingent propositions. One is tempted to say, therefore, that Cleave's system is the most "paradoxical" explication of entailment to be found in the literature.

---

[1] A similar point is made by von Wright ([7], p. 177) and Smiley ([6], p. 239).

The very weakness of the class of rigid tautologies makes one suspect that Cleave might be putting forward a kind of "minimal" theory of entailment—that is, a theory which captures the ground common to all so-far proposed non-classical entailment theories.[1] This suspicion is given some plausibility by the fact that Cleave himself offers his definition 'in view of the absence of agreement on the nature of the paradoxical implications' (p. 118). The only problem with this suggestion is that it is false: for there are Cleavian entailments which are not Anderson and Belnap entailments. Thus, $((p \vee q) \,\&\, \bar{p}) \rightarrow q$ holds in Cleave's system but not in that of Anderson and Belnap (see p. 18 of [1]). However, the set of Cleave entailments *do* form a proper subset of the set of Geach entailments. We prove this as follows: according to Geach, A entails B if and only if either (a) $A \supset B$ is a tautology and neither $\bar{A}$ nor B are tautologies or (b) $A \supset B$ is a substitution instance of $C \supset D$, where $C \supset D$ is a tautology and neither $\bar{C}$ nor D are tautologies.[2] On the other hand, Cleave proves for his notion of entailment (Theorem 5, p. 121) that if $A \rightarrow B$ then neither A nor $\bar{A}$ nor B nor $\bar{B}$ are tautologies. But clearly, for Cleave, if $A \rightarrow B$ then $A \supset B$ is a tautology (in fact, a rigid tautology). Hence, if $A \rightarrow B$ then $A \supset B$ is a tautology and neither A nor $\bar{A}$ nor B nor $\bar{B}$ are tautologies. It follows that if '$A \rightarrow B$' holds in Cleave's system then 'A entails B' holds in Geach's system, since condition (a) above will be satisfied. Thus, all Cleave entailments are Geach entailments. On the other hand, Geach's condition (b) ensures that some Geach entailments will not be Cleave entailments, since Cleave's system is not closed under substitution. Thus, $(p \,\&\, \bar{p}) \supset p$ answers to a Geach entailment since it is a substitution instance of $(p \,\&\, q) \supset p$ and thus satisfies condition (b), while $(p \,\&\, \bar{p}) \supset p$ does not answer to a Cleave entailment since $(p \,\&\, p) \supset p$ is also a tautology. Therefore, the set of Cleave entailments is a proper subset of the set of Geach entailments. So if Cleave's theory is to be thought of as a "minimal" theory, then it can only be considered a minimal version of Geach's theory, and not a minimal version of, say, the Anderson and Belnap system.

Cleave's entailment notion is thus inadequate as an attempt to solve the problem which non-classical entailment theories are intended to solve—that is, to capture (formally) the intuitive notion of logical consequence, or of B's following logically from A, which, it is claimed, strict implication (or the equivalent classical notion of logical consequence) fails to capture. Cleave's system is simply too weak; there aren't enough entailments in it. On the other hand, Cleave's system is contained within the much less restrictive Geachian system. But just how unrestrictive Geach's theory is is difficult to ascertain. It certainly looks more restrictive than strict implication, as can be seen from the fact that transitivity does not hold unrestrictedly in Geach's system whereas strict implication is unrestrictedly transitive. But just *how much* more restrictive Geach's theory is than strict implication is problematic. Smiley, in discussing a system for entailment which is essentially the same as Geach's, writes that the failures of transitivity 'seem to form a fence round the "paradoxes" and to occur nowhere else' ([5], p. 242). Although this claim of Smiley's is, in fact, strictly speaking false,[1] it nevertheless serves to highlight the fact that Geach's theory is one of the least restrictive non-classical entailment theories possible, in that virtually all "non-paradoxical" classical entailments will answer to Geachian entailments although the "paradoxical" ones will not. If even so Geach's theory turns out to be untenable, then this will constitute a strong argument for the identification of entailment with strict implication.

---

[1] I owe this suggestion to Kit Fine, whom I would like to thank for helpful discussion. I am also indebted to David Miller for useful criticism.

[2] I have here slightly altered Geach's definition as it appears in [5]. There Geach spoke in terms of 'theoremhood' in *ES*, the entailment extension of some formal system *S*, rather than in terms of 'tautologousness'. But if we identify *S* with the propositional calculus (*PC*) then, since all tautologies are theorems of *PC* and vice versa, we can replace one by the other. My reason here for giving the altered definition is to make the comparison with Cleave's system more obvious.

The original, intuitive, formulation of the Geach–von Wright theory ([4] and [7]) rests upon the introduction of an epistemic component into the semantics of entailment. Thus, for A to entail B not only must $A \supset B$ be a tautology, but there must exist an *a priori* way of coming to know this which is not simultaneously a way of coming to know the falsity of A or the truth of B. The problem with this way of putting things is that it is unfortunately vague: for while it is true that the truth-table method does not provide a means of coming to know the truth of, say, $(p \& \bar{p}) \supset q$ which is not also a means of coming to know the falsity of $(p \& \bar{p})$, it could certainly be maintained that the Lewis argument itself provides a means of coming to know the truth of $(p \& \bar{p}) \supset q$ which is *not* a means of coming to know the falsity of $(p \& \bar{p})$, for the argument nowhere demonstrates the falsity of $(p \& \bar{p})$.[2] Thus, it is by no means clear that the original formulation excludes the "paradoxes". As a result, Geach was led to introduce (in [5]) a more formal definition, wherein one restricts the available stock of *a priori* methods of 'coming to know' to those available in some *given* formal system S (say, the propositional calculus). Then, A entails B iff either (a) $A \supset B$ is a theorem of S and neither $\bar{A}$ nor B are theorems of S, or (b) $A \supset B$ is a substitution instance of $C \supset D$, where $C \supset D$ satisfies condition (a). Since the Lewis argument is formulable totally within the propositional calculus, the above definition provides a clear-cut means for deciding whether the "paradoxical" conclusion (f) holds; and as can be easily checked, it does not.

---

[1] In [3] A. J. Dale has succeeded in showing that neither $((p \vee q) \& (\overline{\bar{p} \vee q})) \rightarrow (p \& q)$ nor $(\bar{p} \vee \bar{q}) \rightarrow ((p \vee q) \vee (\overline{\bar{p} \vee q}))$ holds in Geach's theory. In other words, the failures of transitivity extend a bit beyond the "pure paradoxical" cases of $(p \& \bar{p}) \rightarrow q$ and $q \rightarrow (p \vee \bar{p})$.

[2] A similar point has been made recently by A. J. Dale in [3], p. 218.

Now Geach's improved theory avoids the "paradoxical" (f) only at the cost of denying the unrestricted transitivity of entailment; for every step in the Lewis argument is sanctioned in Geach's theory except the move from (a)–(e) to (f). But then, so it would seem, the proof-chain (a)–(e) must constitute, for Geach, a *valid* proof-chain. He writes: 'So long as each link in a chain of proof answers to an entailment, the whole chain is sound' ([5], p. 239). So it looks as if, on the basis of $p \,\&\, \bar{p}$, we can validly prove any arbitrary proposition q whatsoever. But this means that if I now *assert* $p \,\&\, \bar{p}$, then I can validly assert any q whatsoever; I can *validly pass* from $p \,\&\, \bar{p}$ to q. But then what have we gained in denying that $p \,\&\, \bar{p}$ *entails* q? Especially since what Lewis himself was concerned with was exactly 'that relation which is present when we "validly" pass from one assertion, or set of assertions, to another assertion' (quoted in [7], p. 166). Seemingly, the only thing which we have gained is that we cannot pass *directly* from $p \,\&\, \bar{p}$ to q (since $(p \,\&\, \bar{p}) \rightarrow q$ does not hold); but since (a)–(e) is a valid proof-chain, it certainly seems that we can so pass *indirectly*. In other words, it begins to look as if Geach's restriction (on transitivity) really comes *too late* in the Lewis argument to avoid the "paradox" successfully. For what logical difference does it make that $p \,\&\, \bar{p}$ does not *entail* q, when we can validly *prove* or *infer* q from $p \,\&\, \bar{p}$? Moreover, remembering that entailment theories are supposed to capture the intuitive notion of B's following logically from A, Geach's theory thus gives rise to the highly "paradoxical" result that although we can validly *prove* q from $p \,\&\, \bar{p}$, it is not the case that q logically *follows* from $p \,\&\, \bar{p}$.

Now the above result is more than just "paradoxical"; it is disastrous. For it follows that the existence of a valid proof of B from A no longer guarantees that A entails B, and therefore that proof is no longer reliable as a method for the *discovery* of entailments. Such a restriction on the reliability of proof is, of course, not too worrying if we can check, independently of proof, whether or not A entails B (as we can in the propositional calculus, *via* truth-tables). But in, say, first-order logic the construction of a derivation is one of our main methods for the discovery of validity, or entailment, and if our proof methods could no longer be relied upon then our ability to discover whether or not A entails B would be severely restricted. And since much of our mathematical thinking takes place in first-order (and higher order) languages, it would follow that the reliability of proof in mathematics would be severely restricted as well. In other words, one would have thought that if one were going to introduce a stronger, non-classical, concept of entailment or consequence then this would be reflected in weaker proof rules, such that if B was derivable from A it *would follow* that A entails B. Yet Geach appears to do the former without doing the latter.

It might, of course, be objected that I have here misunderstood Geach's theory, and that from the fact that some A (say, p & $\bar{p}$) does not entail some B (say, q) it follows that we *cannot* validly prove B from A, and so cannot validly pass from A to B. In other words, it might be argued that, for Geach, we can validly prove B from A *only if* A does, in fact, entail B. Now this *would* put some bite into Geach's restrictions on transitivity, since failures of transitivity of entailment would now extend to proof. Moreover, it would seem to restore reliability (or *soundness*) to proof, since the existence of a valid proof from A to B would now guarantee that A entails B. But the problem with this suggestion is that it would render all proofs of necessary propositions either redundant or circular. For assume that N is some necessary truth (of, say, mathematics) which we want to prove on the basis of some antecedently known mathematical truth A. Assume further that this proof is carried out in a number of steps: i.e., from A to B, B to C, . . . , N–1 to N, where each step answers to an entailment. Can we conclude that we can validly infer N from A? If the proof of N from A is valid (i.e., if transitivity of proof holds) only if A, in fact, entails N, then the answer is clearly 'No'. For we shall first have to know that A entails N *before* we can assert that the constructed proof is in fact a valid proof. But if we can know this independently of the proof, then clearly the proof is redundant. On the other hand, if our only method for discovering whether or not A does, in fact, entail N is *via* the construction of the proof (and this is often the case), then the proof is circular. For we will have to *assume* that A entails N in order that the proof of N from A be a valid proof. In other words, proof is once again valueless as a method of extending our mathematical knowledge. Clearly, the lesson of this argument must be that transitivity is crucial if proof is to do the job we expect of it: for unless proof is unrestrictedly transitive we can't use it to discover entailments. But if proof must be unrestrictedly transitive, and if we introduce, as Geach does, a notion of entailment which is *not* unrestrictedly transitive, then we will be able to prove validly from some propositions conclusions which are *not* entailed by them. Thus, an entailment notion which is not unrestrictedly transitive is unsatisfactory.[1]

The situation might be summed up as follows: either the failure of transitivity in Geach's theory of entailment extends to proof, or it does not. If it does not, then there will be valid proofs (i.e., proofs where each step corresponds to an entailment, as in (a)–(e) of the Lewis argument) such that, although we can validly pass from the premisses to the conclusion, the premisses do not entail the conclusion. But if we can validly pass from *A* to *B*, what is the logical gain in denying that *A* entails *B*? On the other hand, if the failure of transitivity does extend to proof, so that there is a valid proof from *A* to *B* only if *A* does, in fact, entail *B*, then all proofs of (necessary) mathematical truths become either redundant or circular. Hence, either (a) Geach's theory does not *really* avoid the "paradoxes" and it fails to avoid them at the cost of rendering proof unreliable as a guide to entailment; or (b) it does avoid the "paradoxes" but at the cost of rendering mathematical proof either redundant or circular, and so valueless. It follows that the Geachian account of entailment is either, at best, no better off than the classical account, or else it is worse off. Hence, it is worse off.

---

[1] This, incidentally, confirms the strongly held view of Anderson and Belnap (see p. 11 of [1]) that a non-transitive notion of entailment is incoherent.

Finally, remember that the "felt need" for a non-classical entailment theory arises in the following way: if we assume that (a) the truths of mathematics are necessarily true, and if we then (b) identify entailment, or logical consequences, with the classical notions of Lewis or Bolzano and Tarski, then we get the absurd result that (c) all mathematical truths are logically equivalent, or entail each other. Since (c) appears to be clearly false, we must reject either (a) or (b). Thus, in an important sense, the rejection of (b) can be read as a defence of (a). We have argued that, at least in comparison to the theories of Cleave and Geach, the classical account of consequence is to be preferred. Hence we seem to be landed in the "absurd" position of denying (a). But why not? After the discovery of non-Euclidean geometries (a) looks pretty dubious anyway. And after the futile attempts of both logicism and formalism to provide certain foundations for our mathematical knowledge, after the results of Gödel, Cohen, Tarski and others, we no longer need pretend that our mathematical knowledge is infallible. And if our mathematical knowledge is fallible, then what any longer is the *point* in assuming that mathematical truths are necessarily true? But clearly this is not the place to pursue *this* argument further.

University of Edinburgh © LARRY BRISKMAN 1975

REFERENCES

[1] A. R. Anderson and N. D. Belnap, Jr., 'Tautological Entailments', *Phil. Stud.* 13 (1962), 9–24.
[2] John P. Cleave, 'An Account of Entailment Based on Classical Semantics', ANALYSIS 34 (March, 1974), 118–22.
[3] A. J. Dale, 'Geach on Entailment', *Phil. Rev.* LXXXII (1973), 215–19.
[4] P. T. Geach, 'Entailment', *Arist. Soc. Supp.* Vol. XXXII (1958), 157–72.
[5] P. T. Geach, 'Entailment', *Phil. Rev.* LXXIX (1970), 237–9.
[6] T. J. Smiley, 'Entailment and Deducibility', *Proc. Arist. Soc.* LIX (1958–9), 233–54.
[7] G. H. von Wright, *Logical Studies*, RKP, London (1957), 166–91.

ADJUNCTION AND PARADOXICAL DERIVATIONS

By Charles F. Kielkopf

SEVERAL years ago in [1] W. T. Parry presented his sentential logic system of Analytic Implication. The characteristic mark of Parry's system is that an analytic implication formula $(A \rightarrow B)$ is a theorem only if every sentential variable in B is a sentential variable in A. Thus neither $((p \ \& \ {\sim}p) \rightarrow q)$ nor $(q \rightarrow (p \vee {\sim}p))$ are theorems. So the theorems of Parry's system are candidates for relevant implications or entailments. But, as we will see, a system which has a relevance connection between antecedent and consequent in all theorems whose principle connective is $\rightarrow$ may allow derivations in which there is no relevance between premisses and conclusion. In particular, we will see that Parry's system with adjunction tolerates such derivations. Recently, in [2], J. M. Dunn investigated Parry's system. Dunn modified Parry's system in substantial ways as well as in ways which Dunn considered trivial. A substantial modification was addition of $((A \ \& \ {\sim}B) \rightarrow {\sim}(A \rightarrow B))$ as an axiom schema. Dunn considered his use of axiom schemata instead of axioms plus a rule of uniform substitution for variables in theorems a trivial modification. Parry did not explicitly state a rule of adjunction: From A and B, infer $A \ \& \ B$. Dunn also considered his addition of a rule of adjunction as a trivial modification. Dunn wrote 'We hence take the rule of adjunction as a primitive rule of Parry's system, thereby correcting what we take to be a mere oversight on Parry's part.' Presumably Parry himself regarded his failure to present adjunction as a mere oversight. When Parry presented his system some 30 years later on pp. 151–53 of [3], he gives adjunction and *modus ponens* for $\rightarrow$ as its rules. But the addition of adjunction to Parry's system is far from trivial. The main purpose of this note is to show that addition of adjunction to Parry's system allows derivation of q from $(p \ \& \ {\sim}p)$ by use of analytic implications—$(A \rightarrow B)$ which are theorems—and *modus ponens* for $\rightarrow$. Hence, the addition of adjunction to Parry's system prevents it from being a system of logic which protects us from being able to derive an arbitrarily selected claim from an explicit contradiction. Of course, I think that it is a serious defect of Parry's system that it cannot allow adjunction without allowing there to be such a derivation.

I shall now sketch how, with adjunction, q can be derived from $(p \ \& \ {\sim}p)$ in Parry's system. Then I will make a few remarks on the proofs of the theorems used. I shall close by observing that an entailment system of P. Geach cannot allow *modus ponens* and adjunction while a system of J. P. Cleave seems able to accommodate these rules.

(1)	$p \,\&\, {\sim}p$	hypothesis
(2)	$(q \vee {\sim}q)$	theorem
(3)	$(p \,\&\, {\sim}p) \,\&\, (q \vee {\sim}q)$	adjunction (1), (2)
(4)	$(p \,\&\, {\sim}p) \,\&\, (q \vee {\sim}q) \rightarrow (p \vee q)$	theorem
(5)	$(p \vee q)$	*modus ponens* (3), (4)
(6)	$(p \,\&\, {\sim}p) \rightarrow {\sim}p$	theorem
(7)	${\sim}p$	*modus ponens* (1), (6)
(8)	${\sim}p \,\&\, (p \vee q)$	adjunction (7), (5)
(9)	${\sim}p \,\&\, (p \vee q) \rightarrow q$	theorem
(10)	q	*modus ponens* (8), (9)

In Parry's system & and v are commutative and if $(A \rightarrow B)$ and $(B \rightarrow A)$ are theorems B can be substituted for some or all occurrences of A in a formula. In the following remarks I will cite axioms as given in [1]. The crucial axiom for getting (2) is: $(p \rightarrow q) \rightarrow (p \supset q)$. I take Parry's $(p \supset q)$ as $({\sim}p \vee q)$. We have $(q \rightarrow q)$ in Parry's system. So it is easy to get $(q \vee {\sim}q)$. The crucial axiom for getting (9) is: $((p \vee q) \,\&\, {\sim}q) \rightarrow p$. Number (4) above takes a bit more work. An axiom gives distribution of & over v. So we have: $(p \,\&\, {\sim}p) \,\&\, (q \vee {\sim}q) \rightarrow (((p \,\&\, {\sim}p) \,\&\, q) \vee ((p \,\&\, {\sim}p) \,\&\, {\sim}q))$. We have also: $((p \,\&\, {\sim}p) \,\&\, q) \rightarrow q$ and $((p \,\&\, {\sim}p) \,\&\, {\sim}q) \rightarrow p$. We then use adjunction on the last two formulas to get: $(((p \,\&\, {\sim}p) \,\&\, q) \rightarrow q) \,\&\, (((p \,\&\, {\sim}p) \,\&\, {\sim}q) \rightarrow p)$. By use of the axiom: $((p \rightarrow q) \,\&\, (r \rightarrow s)) \rightarrow ((p \vee r) \rightarrow (q \vee s))$ we get: $(((p \,\&\, {\sim}p) \,\&\, q) \vee ((p \,\&\, {\sim}p) \,\&\, {\sim}q)) \rightarrow (p \vee q)$. Because $\rightarrow$ is transitive we get (4). I am not sure that (4) is provable without adjunction.

In [4], P. T. Geach presents a way of selecting a proper subset of the truth-functional tautologies as entailments. For Geach $(A \rightarrow B)$ is a true entailment if and only if $(A \supset B)$ is a substitution instance of truth functional tautology $(A^* \supset B^*)$ such that A^* is not a contradiction and B^* is not a tautology. By Geach's standard $(p \,\&\, {\sim}p) \rightarrow p$ and $(p \,\&\, {\sim}p) \rightarrow {\sim}p$ hold because both $(p \,\&\, {\sim}p) \supset p$ and $(p \,\&\, {\sim}p) \supset {\sim}p$ are obtainable from $(p \,\&\, q) \supset p$. But $(p \,\&\, {\sim}p) \rightarrow q$ fails because any tautology of which $(p \,\&\, {\sim}p) \supset q$ is a substitution instance will have a contradictory antecedent. Geach does not give any rules for deriving some formulas from others. His system is simply a subset of the truth-functional tautologies. But if Geach were to give *modus ponens* for $\rightarrow$ and adjunction as derivation rules, q could be derived from $(p \,\&\, {\sim}p)$ as below.

(1)	$p \,\&\, {\sim}p$	hypothesis
(2)	$(p \,\&\, {\sim}p) \rightarrow p$	true entailment from $(p \,\&\, q) \supset p$
(3)	$(p \,\&\, {\sim}p) \rightarrow {\sim}p$	true entailment from $(p \,\&\, q) \supset q$
(4)	p	*modus ponens* (1), (2)
(5)	${\sim}p$	*modus ponens* (1), (3)
(6)	$p \rightarrow (p \vee q)$	true entailment $p \supset (p \vee q)$ is a suitable tautology

(7) $(p \vee q)$ *modus ponens* (4), (6)
(8) $\sim p \,\&\, (p \vee q)$ adjunction (5), (7)
(9) $\sim p \,\&\, (p \vee q) \to q$ true entailment from $\sim p \,\&\, (p \vee q) \supset q$
(10) q *modus ponens* (8), (9)

I do think that it is a defect in Geach's system that it cannot be supplemented with these reasonable derivation rules without tolerating the above derivation.

In closing, it is interesting to observe that J. P. Cleave's system of [5] could be supplemented with *modus ponens* and adjunction without tolerating the above derivation. Cleave's system is also a proper subset of the truth-functional tautologies. For Cleave $(A \to B)$ is a true entailment if and only $(A \supset B)$ is a truth-functional tautology which has no subformula F such that when F is replaced in $(A \supset B)$ with $\sim F$ to get $(A \supset B)^*$, $(A \supset B)^*$ is not a truth-functional tautology. Thus for Cleave neither (2) nor (3) of the immediately preceding derivation are true entailments. Both $(p \,\&\, \sim p) \supset p$ and $(p \,\&\, \sim p) \supset \sim p$ remain tautologies if the rightmost occurrence of p is replaced with $\sim p$. I am not here defending Cleave's system; I am simply noting that on one standard it is better than Parry's or Geach's.

The Ohio State University

© CHARLES F. KIELKOPF 1975

REFERENCES

[1] Parry, W.T. 'Ein Axiomensystem für eine neue Art von Implikation (analytische Implikation)', *Ergebnisse eines mathematischen Kolloquiums*, vol. 4 (1933), pp. 5–6.

[2] Dunn, J.M. 'Parry's Analytic Implication', *Notre Dame Journal of Formal Logic*, vol. 13 (1972), pp. 195–205.

[3] Parry, W.T. 'Logic of C. I. Lewis', pp. 115–154 of *The Philosophy of C. I. Lewis*, ed. P. A. Schilpp, Open Court Pub. Co., La Salle, Illinois, 1968.

[4] Geach, P.T. 'Entailment Again', *The Philosophical Review*, 79 (1970), pp. 237–9.

[5] Cleave, J.P. 'An Account of Entailment Based on Classical Semantics', ANALYSIS 34.4, pp. 118–22.

LOGICAL SUBTRACTION

By James L. Hudson

In a recent article Robert A. Jaeger has raised doubts about the meaning of the interrogative sentence: 'What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?'[1] To utter this sentence with the intention of asking a question is apparently to presuppose that we know what "subtraction" in the logical sense is; but how, if pressed, could we define this notion? Jaeger examines several possible definitions, fails to find a suitable one, and concludes or at least suggests that such apparent questions as the one quoted above are only meaningless pseudo-questions. I shall argue, on the contrary, that one of the definitions which Jaeger rejects is actually correct, and that questions and statements framed in terms of logical subtraction are unexceptionable.

Jaeger's method in seeking a definition, at least initially, is to rely on the analogy between logical subtraction and the physical operation of taking away, as when some stones are taken away (= subtracted) from a pile of stones. In other words, the subtracting of one proposition from another is to be conceived by analogy with the taking away of a subpile of stones from a given pile. Jaeger mentions another possible analogue —arithmetical subtraction—but he makes virtually no use of it, and indeed it would be of no help to him (see n. 3 below).

On the basis of his analogy Jaeger suggests the following rules for the operation of logical subtraction. Let P, Q and R be statements (facts or falsehoods, propositions). R is what is left over when Q is subtracted from P (symbolically, $R = P - Q$) only if (a) P implies R ('What is left over will be a part of the original whole'); (b) the conjunction of Q and R (symbolically, 'Q & R') implies P (or, better, is equivalent to P; 'The whole is equal to the sum of its parts'); (c) R does not imply Q ('What is subtracted cannot be a part of what is left over'); (d) Q does not imply R ('What is left over cannot be a part of what is subtracted'); (e) given P and Q, R is uniquely determined—i.e., subtraction is a proper function ('When a part is subtracted from a whole, there is exactly one part left over').[2] Jaeger states in (b) above only that Q & R *implies* P. But he elsewhere assumes that Q & R *is equivalent to* P, as he is entitled to do by the slogan 'The whole is equal to the sum of its parts', and this shall be my interpretation of (b). I will write this 'Q & R $= P$', on the theory that distinct propositions cannot be logically equivalent to each other. I think that this theory is correct, but anyone who doubts it can read '='

---

[1] Robert A. Jaeger, 'Action and Subtraction', *The Philosophical Review*, LXXXII (1973), 320–29. The sentence is from Wittgenstein's *Philosophical Investigations*, part I, para. 621.

[2] Jaeger, 320.

as 'is equivalent to' rather than as 'is identical with'; or, better, he can take 'P', 'Q', etc., to refer to equivalence classes of propositions rather than to individual propositions, and can make the corresponding adjustment in his interpretation of '&', '$-$', etc.

The quotations in parentheses above are given by Jaeger evidently because they express principles which we immediately recognize as valid for physical subtraction, and hence as justifying the analogous principles for logical subtraction.[1] In fact, however, the physical principles are valid only if their application is understood to be somewhat restricted. The corresponding restriction for logical subtraction would be to cases in which P implies Q but Q does not imply P. I believe Jaeger would accept this restriction, for in framing these principles he seems to have had one eye on his example about raising one's arm, in which indeed the fact that I raise my arm implies but is not implied by the fact that my arm goes up. In any case, let me impose this restriction for the present.

Not surprisingly, the first candidate Jaeger considers for the role of $P - Q$ is the material conditional $Q \supset P$. This satisfies (a)–(d) above; but, as Jaeger points out, it is not the only proposition that does so. $(Q \vee S) \supset P$ for any statement S will also satisfy (a)–(d); and in general $(Q \vee S) \supset P \neq Q \supset P$. Jaeger apparently regards this circumstance as counting strongly against $Q \supset P$'s candidacy; but all it really shows is that (a)–(d) are not in themselves sufficient to pick out $P - Q$ uniquely —we need to add some further condition. Now among the propositions which satisfy (a)–(d) a special place is occupied by $Q \supset P$, in that all the others imply it. (The Deduction Theorem assures us that if Q & R implies P, then R implies $Q \supset P$.) And it seems reasonable to stipulate that if there are several different propositions whose conjunction with Q is P, then the least (i.e., the weakest) of these shall be considered *the* difference between P and Q. This additional stipulation, then, along with the aforestated (a)–(d), gives us a unique determination of $P - Q$: it determines that logical subtraction be the converse of material conditionalization.

Among the reasons given or hinted at by Jaeger for rejecting this conclusion there is one which appears to me to deserve discussion here. Before taking it up, however, I must digress briefly upon the topic of logical independence.

To say that two propositions are logically independent is to say that certain logical relations fail to obtain between them. As a minimum we should require that neither proposition imply the other; this weak notion of independence is used when we speak of the independence of axioms in an axiomatic system. We might require as well that the propositions should not be contraries (i.e., that neither imply the negation of the other), or that they should not be subcontraries (i.e., that neither be implied by the negation of the other), or that they be neither contraries nor subcontraries. This last and strongest notion of independence has the greatest currency in philosophical discussions. For example, it is this notion that the logical atomists had in mind when they stipulated that any two atomic facts must be "independent". And if we believed in a distinction of contingent facts into physical and psychological, we would presumably maintain that any physical fact is (strongly) independent of any psychological fact.

---

[1] The fact that (c) and (d) seem to be reasonable requirements to place on logical subtraction shows that the latter is conceived better by analogy with physical subtraction than by analogy with arithmetical subtraction. For in arithmetic there is nothing corresponding to (c) and (d); indeed the number taken away and the number remaining might be identical, as in '$8 - 4 = 4$'.

According to my proposed definition of subtraction (still under the restriction that P implies but is not implied by Q), Q and $P-Q$ are independent in the weak sense; but they *are* subcontraries (for either follows from the negation of the other), and so they fail a stricter test for independence. Now, do we have intuitive grounds for requiring this stricter independence of Q and $P-Q$? Jaeger apparently thinks so, and I hypothesize that his opinion is based in part on an analogy with physical subtraction; for each stone in a pile is quite independent of each other stone, and disjoint subpiles are independent of each other. Thus a logical atomist, thinking of the atomic facts Q and R as being like distinct stones and of the molecular fact Q & R as like a two-stone pile, might naturally want $(Q \,\&\, \mathrm{R})-Q$ to be R rather than $Q \supset (Q \,\&\, \mathrm{R})$ $[= Q \supset \mathrm{R}]$. And a believer in the physical-psychological distinction might have the same desire in cases (such as Jaeger's arm-raising example) in which he regards Q as a purely physical fact and R as a purely psychological one.

But to reason thus is to forget the very analogy between propositions and subpiles of stones on which we have been relying. According to this analogy, what corresponds to strong logical independence between two propositions is the failure of either of two subpiles of stones to be wholly contained in *or wholly excluded from* the other. Thus two subpiles would have to overlap somewhat (i.e., to have a stone or stones in common) in order to correspond to two strongly independent propositions. On the other hand, two subpiles which did not overlap at all, and which might thus be said to be "physically independent", would correspond to two propositions which, while weakly independent, were subcontraries. By the same token Jaeger is wide of the mark when he contrasts the fact that physically independent piles 'have no parts in common' with the fact that any two propositions Q and R must have a 'part in common', *viz.* $Q \vee \mathrm{R}$ (Jaeger, 329). For according to our analogy Q and R will still correspond to the independent piles of stones provided $Q \vee \mathrm{R} = S \vee \sim S$ (which counts as "nothing"), and this is the case precisely if Q and R are weakly independent subcontraries.

Furthermore there is a severe difficulty in requiring that Q and $P-Q$ be strongly independent while also insisting on (a)-(d): namely, that no unique general determination of $P-Q$ will be possible. For $P-Q$ will have to imply $Q \supset P$ and be implied by P without being equivalent to either. But there is no one proposition which satisfies these requirements, nor is there any plausible addition to (a)–(d) which would narrow the field down to one. The point can be illustrated with one of Jaeger's examples. He writes: ' "X is red" properly implies "X is coloured". Is there a third statement which has all the required features? What is it?' (325). We would have said that 'X is coloured $\supset X$ is red' was the required difference between 'X is red' and 'X is coloured', except that it is only weakly independent of 'X is coloured'; what we need in order to get strong independence is a statement which properly implies 'X is coloured' and is properly implied by 'X is red'. '(X is coloured v X is spherical) $\supset X$ is red' fills the bill; but so does '(X is coloured v X weighs a ton) $\supset X$ is red', etc. In short, more than one statement lies between 'X is coloured $\supset X$ is red' and 'X is red'; and in general more than one statement will lie between $Q \supset P$ and P. Under the influence of a special interest we may have a natural way of picking out one of these statements—we may specify the one that is atomic, or the (weakest) one that is purely psychological. But there is no *general* condition which we can add to (a)–(d) which will do the job of selecting a single statement in a way appropriate to the term 'subtraction'. So given that subtraction is an "operation" in the mathematical sense, Q and $P-Q$ can be required at most to be weakly independent; and we are left free to argue as I have done above that logical subtraction is the converse of material conditionalization.

The reasons which led Jaeger to reject this conclusion should have led him instead to rephrase his problem about arm-raising so as not to use a term for subtraction. For his concern with strong independence shows that he has slid from the question of volition as he originally stated it to a slightly different question which can be formulated thus: 'What is the weakest purely psychological statement whose conjunction with "My arm goes up" implies "I raise my arm"?' Although there is no one statement—even no one psychological statement—whose conjunction with 'My arm goes up' implies 'I raise my arm', there is sure to be a weakest one (one that all the others imply). However, this (call it 'R') cannot be considered *the difference,* since we have specified that R be "psychological", and this clearly is not part of the meaning of 'subtract'. Jaeger's *new* problem about arm-raising does indeed require the strong independence of R from 'My arm goes up'; but it has nothing to do with subtraction.

So far the discussion has been restricted to cases of subtraction in which P (the minuend) implies Q (the subtrahend) but not conversely. To these correspond the very cases involving a pile of stones in which it is most common to speak of "subtraction"; namely, those in which the stones taken away comprise some but not all of the stones originally in the pile, and in which furthermore all the stones said to be "taken away" did originally belong to the pile. We would like now to extend the definition of logical subtraction by removing this restriction, so that $P-Q$ will be well-defined for *any* pair of propositions P and Q. Perhaps no one will object if we simply drop the restriction and let $P-Q = Q \supset P$ for all P and Q. But we can find a bit of support for this procedure, beyond its inherent attractiveness, by noting that a similar extension has already been made in set theory.

Set-theoretical subtraction, as its name suggests, is an operation on sets which is modelled after physical subtraction. And it is clear from the physical analogue that for any sets X and Y, if Y is a non-empty proper subset of X then $X-Y$ must be the set of elements of X which are not elements of Y; i.e.:

(1) $X-Y = X \cap \bar{Y}$.

It is not so immediately clear how to understand subtraction when the subtrahend is not a non-empty proper subset of the minuend; but it has been found that the most natural extension of the notion is simply to adhere to (1) in all cases. Thus we have in (1) a definition of set-theoretical subtraction which is unrestricted within the field of sets.

Now, we have concluded from our previous discussion of logical subtraction that $P-Q = Q \supset P$, i.e., that $P-Q = P \vee {\sim} Q$, in the restricted range of cases. Comparing this with (1) we find an analogy between the two kinds of subtraction, in which set-theoretical intersection corresponds to logical disjunction and set-theoretical complementation to logical negation. (This is the dual of the usual correspondence, in which intersection is linked with conjunction rather than with disjunction.) The naturalness of this analogy can be checked *a posteriori;* for example, it makes the empty set correspond to the necessary proposition, and surely to "take away nothing" is in the set-theoretical case to subtract the empty set, while in the propositional case it is to subtract $S \vee {\sim} S$, which "says nothing". So this analogy, which is in effect an indirect analogy between logical and *physical* subtraction, gives us an additional warrant for regarding $P-Q$ to be $Q \supset P$ in all cases.

By treating logical subtraction as an operation on propositions generally, we force ourselves to give up some of Jaeger's principles listed above. We have to give up the strong form of (b) and revert to the weak form; for Q & $(P-Q)$ will not be equivalent to (identical with) P unless P implies Q, though it will in any case imply P. By allowing the subtrahend to be the necessary proposition we relinquish (c), and by permitting the minuend and the subtrahend to be identical we surrender (d). Our operation still deserves to be called 'subtraction', however, for all these remarks about logical subtraction have set-theoretical analogues, and indeed (though we do not ordinarily use the term 'subtraction' for the physical operation except under the previous restriction) they have physical analogues as well.

I conclude that the converse of material conditionalization is the one and only logical notion which deserves the name 'subtraction'. Among logical operations it bears the closest analogy to the physical operation of that name, and it is isomorphic with set-theoretical subtraction (as these two operations appear in propositional logic and in the calculus of classes, respectively). My conclusion does not in itself solve any important traditional problem of philosophy; it does not, for instance, throw much new light on the philosophy of mind. But it does show that talk of "subtracting" one statement from another has a perfectly natural meaning, and thus it saves Jaeger's problem of volition (as originally stated) from being "dissolved".[1]

Northern Illinois University © James L. Hudson 1975

CONFUSING THE AUDIENCE

By G. H. Bird

In most of the recent work on speech acts a convenient device is used to simplify references to the speaker (S) and the audience (A) to whom the utterance is addressed. In Searle's *Speech Acts,* for example, references to speaker and audience (designated merely as a hearer, H) play an important part in the constitutive rules governing speech acts such as promising. In Schiffer's *Meaning*, too, the analysis is generally conducted in such terms, although there is also a careful consideration of audienceless cases.[2] Strawson, in his paper on intention and convention in speech acts, hints at one difficulty in identifying an audience to whom certain convention-bound utterances are addressed, but he does not elaborate the point.[1]

---

[1] The results of this paper could be used to show what precise answers would be like to such questions as: 'What is the difference between rationalism and empiricism?' 'What is the difference between dogs and cats?' 'What is the difference between Johnson and Goldwater?' But in asking a question of this kind we ordinarily do not require or expect a precise answer; we just want a short list of important properties possessed by the one thing and not by the other. In contrast, the philosophical context of Jaeger's Wittgensteinian question justifies him (and me) in trying to answer it precisely. In fact Wittgenstein's intention in introducing the word 'subtract' here was apparently to put us in mind of physical and arithmetical subtraction, and thus to contrast definite problems framed in those terms with the indefinite problem about volition. He wanted to suggest, as Jaeger says, that the latter problem, strictly construed, is unsolvable. I have argued otherwise.

[2] S. Schiffer, *Meaning* (OUP, 1972). See especially pp. 73, ff.

The suggestion I wish to make is that this device is in some respects over-simple and misleading. It is, of course, obviously true that the simple one to one conversation piece is merely one standard case of communication, and that often the audience consists of more than one person. But that by itself would be of no great importance so long as the members of the audience all shared roughly the same relation to the speaker and to his speech act. It is not, then, mere numerical diversity in the audience which complicates the picture, so much as functional diversity.

The simplest way to bring this out is perhaps to follow up Mr. Carter's examination of Searle's condition 4 for the speech act of promising.[2] Mr. Carter has offered some apparent counter-examples to condition 4, on the supposition that that requirement is unambiguous, but I shall suggest that his assumption is itself mistaken. The point, however, does not count against Carter's general thesis but rather supports it, for the ambiguities in condition 4 provide further apparent counter-examples to Searle's thesis as it stands.

Condition 4 in Searle's analysis, a necessary condition of 'happy promising', states that the hearer, H, would prefer the speaker's (S's) doing the promised act to his (S's) not doing it, and further that S believes that H would have that preference.[3] But the relevant ambiguity exists already in the analysandum, formulated as:

> Given that a speaker, S, utters a sentence, T, in the presence of a hearer, H, S sincerely and non-defectively promises that *p* to H (*SA*, p. 57).

Carter has already noted one obscurity in Armstrong's formulation, namely that the audience is to be not just anyone who 'happens to overhear the promisor', but a person *to whom* the promise is made. Searle's analysandum, although expressed in terms merely of a hearer, evidently avoids that difficulty by setting out explicitly to analyse S's promising that *p* to H. But the notion of promising something *to* somebody is itself ambiguous.

Consider the following four cases:

(1) S promises that he will buy his son a bicyle for his next birthday (*p*) in the presence of A_1 who overhears him

(2) S promises that *p* in the presence of a neutral witness A_2 (promises that *p to* A_2)

---

[1] P. F. Strawson, 'Intention and Convention in Speech acts', p. 35 in Searle (ed.) *The Philosophy of Language* (OUP, 1971), where he says: It is not even possible, in other than a formal sense, to isolate among all the participants in the procedure (trial, marriage, game) to which the utterance belongs, a particular audience to whom the utterance can be said to be addressed.

[2] W. R. Carter, ANALYSIS 33.3, pp. 88–92.

[3] Searle, *Speech Acts* (CUP, 1969), p. 58. There is a residual unclarity in Searle's use of 'would' in the condition, rather than simply 'does'; but I do not think this affects my argument.

(3) S promises to his wife, A_3, that *p*, in response to her pressure

(4) S promises to his son, A_4, that *p*, in response to his pressure.

Now it would clearly be wrong to say that S had promised anything to A_1, and it may be thought that this is also true of cases (2) and (3). Yet this itself would surely be wrong, for in some sense S has promised something to A_2, A_3 and A_4; indeed in one way he has promised exactly the same thing to each of them, i.e. he has made the same promise *to* each of them. But it is equally obvious that 'promising to' is not necessarily the same relation in each of the three cases.

First there is a plain distinction between (2) and (3) on one side and (4) on the other. In (2) and (3) the promise is *addressed* to audiences who are not the principal beneficiaries of the promise. In this they are unlike (1), in which the promise is not addressed to A_1 at all, and unlike (4), in which the promise is addressed to what I shall call the 'principal beneficiary'. It may be said that it is (4) alone which Searle's condition is intended to analyse, but as far as the locution 'S promises that *p* to H' goes it certainly also covers cases (2) and (3) as well.

Second there is also a division between cases (2) and (3), which however, may seem either less important or more dubious. For the relation that exists between S and A_2, although common enough in legal contexts, may seem somewhat artificial in the context of promising. It nevertheless seems to me to have a place in promising, and anyway serves to bring out a complexity, which might otherwise pass unnoticed, in some cases like (3). For the natural division between (2) and (3) is between the case of a neutral observer, a witness, and that of an involved or interested party. One standard context for type (3) cases would be that in which S makes the promise to A_3, because A_3 is directly interested both in exacting the promise and in its subsequent fulfilment. But it may, no doubt, happen sometimes that wives are the recipients of such promises, when in reality they do not have that kind of interest, and where S's belief that they do is actually mistaken. In such cases the relation of the audience to S will approximate more closely to that of A_2 to S. It is worth noting that in so far as that interest is found in type (3) cases the line between (3) and (4) is to some extent blurred. For if A_3 is in that relation to S, then she is to some extent a beneficiary of the promise also. But the similarity between (3) and (4) should be a motive for keeping them separate rather than overlooking the differences between them. For the benefits which A_3 naturally expects from the fulfilment of the promise are certainly different from those expected by A_4. In Searle's account of promising, in which he says 'a promise is a pledge to do something *for* you' (*SA*, p. 58), it is clear that cases (3) and (4) are simply conflated. In (3) S certainly gives a pledge to do something for his wife, and in so doing also gives a pledge to do something for his son. But merely to say that would be to overlook the crucial differences between the two cases.

It would no doubt be helpful to stipulate the boundaries between the distinguished cases, and also provide some vocabulary to label them. With some qualms then I suggest that A_1 be labelled a hearer, A_2 an audience, A_3 an addressee and A_4 a principal beneficiary, where these are related in the following ways. It is sufficient to be an A_1 that a person merely hears what was uttered, but that is not sufficient though it is necessary for someone to be an A_2. Being an A_2 is not, similarly, sufficient for someone to be an A_3, but it is a necessary condition for being an A_3. However, finally, to be an A_4 it is neither necessary nor sufficient to be either an A_1 or an A_2 or an A_3. Thus in such a scheme A_1, A_2 and A_3 are defined in terms of speech act situations, and do not exclude each other. A_4, on the other hand, is not defined in terms of such situations at all, though, of course, it may happen that an A_4 is also an A_1, an A_2 and an A_3.

Certainly there are many other ways in which the basic distinctions might be captured in a general scheme, and I claim only that this scheme usefully clarifies the distinctions, not that it is necessarily the best way of bringing them out. One peripheral benefit is perhaps that in separating A_4 from A_1–A_3 it offers some explanation for the odd conflict that seems to have arisen in this area between speech act theorists and moral philosophers.[1] For the scheme brings out a natural division of interest between the moral philosopher and the speech act theorist. The former is naturally interested in the relation between promisor and principal beneficiary, whether the latter happens to be the audience of a speech act or not. Hence, although the term 'promisee', as a person *to* whom a promise is made, shares in the ambiguity of the latter expression, it has most often been intended in moral philosophy to refer to A_4, that is, to the principal beneficiary. But speech act theorists naturally have an interest in the relation between speaker and audience in speech acts, and this is a quite different kind of relationship. Even where the speech act is that of promising it does not follow that the audience is a promisee in the A sense at all.

However, it also seems clear that a failure to separate these relationships has led to ambiguity in the use of the term 'promisee' (which I have therefore deliberately avoided) and to mistakes about promising. Consider, for example, two passages from Mr. Hanfling's recent paper:

> The point is rather that the person to whom the promise is addressed must have an interest in whatever it is that is being promised (p. 15).
>
> It is not a promise if the listener has no interest in what one is going to do; nor can one be said to 'break a promise' if that is the case (p. 25).

It should now be clear that in these two quotations the 'listener' (A_1) and the 'addressee' (A_2 or A_3) are being confused with a beneficiary of kind

---

[1] See, e.g., O. Hanfling, 'Promises, Games and Institutions', *PAS*, 1974.

A_3 or A_4. What is claimed is not, for example, true of a hearer (A_1) or of an audience (A_2) in my senses; and if it is true of an addressee (A_3) or a principal beneficiary (A_4) in my senses, it is still not true of each of them in just the same way.

Just the same objections apply to Searle's condition 4, which differs essentially from Hanfling's claim only in its use of the term 'preference' instead of 'interest'. If it is appropriate to speak of A_2, A_3 and A_4 as people to whom promises are made, then Searle's condition will hold, but in different ways, for A_3 and A_4, but it will not hold at all for A_2. It might be claimed, of course, that the A_2 cases are not genuine examples of promising at all, so I offer the following case for consideration. A husband promises to his wife that he will buy his son a birthday present, in the (mistaken) belief that she wishes the child to have such a present. Actually the wife, let us suppose, has no such preferences, but still takes the utterance as a commitment to the relevant course of action. Here one might perhaps raise questions about the origins of such a promise, but I cannot see any good reason to deny that a promise was made (to the wife) and would be taken by all the parties concerned to have been made. Nor does there seem any good reason to deny that the wife might report to her son that such a promise had been given to her.

If these views are correct, then apart perhaps from the avoidance of unnecessary confusions in talking generally about promising, their main importance will be for the speech act theorist rather than for the moral philosopher. It may therefore be worth noting two general places at which such complications in the audience may be relevant to the discussion of speech acts. First it may help to clarify the asymmetry between speaker and audience entailed by recent accounts of speech acts. For under the influence of Austin's classification of illocutionary forces, and Grice's intentional account of speaker-meaning, the speaker has been given a priority in the determination of such acts. One qualification to this has been the persistent requirement, variously explained, that the audience must understand the utterance, but this requirement itself needs further elucidation, which a more subtle treatment of the audience may prompt. Second the point may help in the classification of speech acts. For certainly there is a prima facie division to be drawn between those acts where the principal beneficiary need not be part of the audience and those acts where he must be. 'Complimenting', for example, would be a case of the former, and 'flattering' a case of the latter. I can compliment Jones (A_4) in his absence to an audience (A_2 or A_3); but it seems that I cannot flatter Jones in his absence, even though the audience may also contain other people to whom the utterance is addressed.[1]

Stirling University © G. H. Bird 1975

---

[1] This distinction was pointed out to me in a general discussion of the differences between complimenting and flattering by Mr. J. C. Boal.

ARE "CAMBRIDGE" CHANGES NON-EVENTS?

By Paul Helm

AT the end of his paper 'Noncausal connections' (4) Jaegwon Kim raises the question of whether it is useful to treat what he calls ' "Cambridge" events' (4(48)), e.g., someone's becoming a widow, as an event. Professor Kim argues that it is, and that such events stand in a relation of non-causal dependence to other events. The event which is Xanthippe's becoming a widow, for example, stands in a relation of non-causal dependence to Socrates' death. In disputing this I shall not be concerned with his remarks about 'agency dependence' (4(44f)), or 'compositional dependence' (4(51f)), but only with the question of whether what he calls ' "Cambridge" events' are cases of non-causal dependence. I shall argue that "Cambridge" events are not events, and *a fortiori* cannot stand in a relation of dependence to other events, whether of causal or non-causal dependence.

Professor Kim takes the idea of "Cambridge" events from Geach (2(71)), who argues that the only sharp criterion that we have for a thing's having changed is the "Cambridge" criterion, that of a change in what can truly be said of an individual. Geach goes on to point out that this is unsatisfactory as a criterion because it allows us to count as changes what, intuitively, are not changes at all. Socrates' being admired by one more person on Tuesday than he was on Monday implies that there are things true of Socrates on Tuesday that are not true of him on Monday. But we are reluctant to say that Socrates has changed from Monday to Tuesday. The "Cambridge" criterion of change is clearly a necessary condition for the occurrence of those events that we regard as being real, such as Socrates' dying, but we shall be concerned only with those events for whose occurrence the "Cambridge" criterion is necessary and sufficient, and not merely necessary. We shall refer to these as *merely* "Cambridge" changes or events.

To begin with, I wish to argue that merely "Cambridge" events are not real events, on the following grounds.

1. Merely "Cambridge" events can "happen" to things that do not exist. But whatever events are, they happen to real individuals, and not to possible but unactual individuals. They are part of the life-history of individuals.

Not only does Xanthippe not need to be married to Socrates in order for her to be the subject of merely "Cambridge" changes upon his death, she need not even exist. Suppose she had pre-deceased him. Then we could say, on the day of Socrates' death, 'Had she been living, and married to Socrates, she would have been widowed today'. Before the death of Socrates it would have been false that had she been married to him she would have been his widow. But after his death it becomes true that had Xanthippe been married to him she would have been his widow.

Another kind of case is that mentioned by Kenny (3(268)), in which I can, by begetting a child, make my long-dead father into a grandfather. My father no longer exists, but certain relational propositions become true of him that were not true of him previously. But these cannot properly be described as events. Events happen to things, but they cannot happen to non-existent things. If a merely "Cambridge" event can happen to a non-existent thing a merely "Cambridge" event cannot be an event.

Similarly with the objects of intentional verbs such as 'believe' and 'talk about'. Smith can come to believe in God, and so it follows that there is something true of God now that was not true of him previously, that he has the property of being believed in by Smith. And likewise Cerberus can become the most talked-of beast of fable. Coming to believe in God, and coming to talk of Cerberus, can both be said to be events. But coming to be believed in, or to be talked about, are not events, because they are compatible with the non-existence of the objects of belief, or of what is talked about.

2. It might be argued that Kim's criterion of merely "Cambridge" change does not allow cases where the individuals undergoing such changes do not exist. This may be so, although he takes his account of such changes from Geach, and *his* account allows this (2(72)). Even if we suppose, on an amended account of merely "Cambridge" change, that merely "Cambridge" changes can only happen to what exists, supposing such changes to be *events* of which it makes sense to ask what their relation of dependence to other events is has curious consequences.

Socrates' dying is certainly an event, and, since it is logically impossible to remain the wife of one's dead husband, Xanthippe's status changes. But then so does the 'status' of everyone related to Xanthippe. For example, Xanthippe's mother becomes the mother of a widow, her sister the sister of a widow. It is purely accidental that there is not one word in English for the phrase 'widow's mother' or 'widow's sister'. Further, suppose that Xanthippe was not in fact married to Socrates. Then, although she did not become a widow on his death she *would have become* a widow upon his death had she been married to him. That is, on what Kim calls the merely "Cambridge" criterion of change, under these circumstances Xanthippe changes, for there is something true of her after Socrates' death that was not true of her previously. What is true after his death is that she would have been a widow had she been his wife. The idea that someone who said, upon the death of Socrates, of Xanthippe (whom (we suppose) was not married to Socrates) 'She would have become a widow had she been married to Socrates' would have been describing something that happened to Xanthippe seems a little far-fetched.

Why does the idea of the possibility of Xanthippe's having become a widow being an event seem far-fetched? It expresses something that is true of Xanthippe, but was not true previously, and so is a merely "Cambridge" event. This is far-fetched, because however difficult it may be to give a satisfactory account of real events or changes, we want to say, intuitively, that events happen to things, that they happen at particular places and take up so much time. The intuitive contrast we wish to draw, however difficult it is to give a proper account of it, is between what happens to something and what does not happen to that thing. But the idea that 'becoming a widow if married to Socrates' describes one of the events in each of the lives of the women who could have been married to Socrates only succeeds in collapsing events into non-events.

There are many relations in which the dying of Socrates stands, perhaps an infinite number. For example, before Socrates' death we may suppose that there was an individual (let us call him 'Alexander') who was truly describable as the latest man to have died in Athens. Upon Socrates' death, Alexander loses this honour, for Socrates becomes the latest man to die in Athens, until the next death. And perhaps there is some relation R such that the death of Socrates stands in R to everything. Or if not one relation, a disjunction of relations. If we suppose that Socrates' death has a unique place and time, this is obviously true. Thus when Socrates dies, events happen to everything. And since something is happening at every moment everything (actual and non-actual) is changing at every moment.

If we come to think that some other condition besides the "Cambridge" criterion is needed for event-hood, then it becomes misleading to ask questions about Xanthippe's becoming a widow *as though* it were an event. The proper conclusion to draw is not that the event of Xanthippe's becoming a widow is an event that is not caused by anything, standing in a relation of non-causal dependence to other events, but that it is not an event at all. The fact that, in the example, Xanthippe is in existence at the time of Socrates death, though relevant to the question of whether or not she *becomes* a widow (as opposed to *would have become* a widow), is not at all relevant to the question of whether or not she undergoes a merely "Cambridge" change. So to suppose that such facts as that the spatial location of Xanthippe's becoming a widow differs from that of Socrates dying has something to do with deciding whether Socrates dying is the *same event as* Xanthippe's becoming a widow (4(42)) is to be misled by a special case. Suppose my father is living. His spatial location when he becomes a grandfather, his proximity to the birth of my son, has nothing to do with whether his becoming a grandfather is a separate event from my wife's becoming a mother or my becoming a father.

The fact of the non-reversibility of the relation of Socrates' death to Xanthippe's becoming a widow ought not to bother us. It is true that

26 OCT 1975

Xanthippe's widowhood depends on Socrates' death, and not *vice versa.* This, however, is not because we here have two events standing in a relation of asymmetrical non-causal dependence, for as we have seen we have no reason to suppose that these are two events. Nor is it because 'being the wife of' is an asymmetrical relation. It is because one description 'the death of Socrates' describes an action or event, and the other 'the widowing of Xanthippe' describes the logical consequence of the action or event, given a certain relation between Socrates and Xanthippe. A widow (or widower) is something someone can become only by something really happening to someone else. The logical consequence of the event of Socrates' dying is not another event.

The word 'consequence' must not mislead us here. The causal consequence of E_1 (if it has causal consequences) is another event (or events) E_2. But the logical consequence of a true description of E_1 is just a further true description of E_1. Not something that happens as a result of E_1's happening, but an implication of a true description of E_1. It is a logical consequence of the action properly described as Jones' having opened the window that the window is open. That is to say, if 'Jones has opened the window' is a true description of what went on, then 'the window is open' is also a true description. But it is not a description of another event, it is part of an extended description of the one event of Jones' having opened the window. In the same way someone who says that Caesar's death is a logical consequence of Brutus's fatal stabbing of Caesar means that the total event properly described as Brutus fatally stabbing Caesar includes Caesar's death as a part of it.

Given the husband-wife relationship between Socrates and Xanthippe it follows logically from Socrates' death that Xanthippe is a widow, and that Xanthippe's mother is the mother of a widow, and so on. And so, as described, there *could* not be a relation of mutual dependence between Socrates' death and Xanthippe's widowhood. There is no question here of a 'counterfactual dependency' (4(44)) of one event upon another. There is an irreversibility, but it is more like the sort of irreversibility that exists between 'All men are mortal' and 'Socrates is mortal' than the irreversibility of a causal relationship.

What this shows is that there could not be a world where the only changes were merely "Cambridge" changes. Merely "Cambridge" changes are parasitic on those changes for which the "Cambridge" criterion is not a sufficient condition of change.

On one view of the analysis of action, there is a special case to be considered here. If the hemlock was administered to Socrates by Xanthippe herself, then, on certain views of the identity of action Xanthippe's killing Socrates is the same action as Socrates' dying which in turn is the same action as Xanthippe's widowing herself. In this case there is a relation of reversibility or symmetry, the symmetry of identity.

Socrates' dying *is* Xanthippe's widowing herself, and Xanthippe's widowing herself *is* Socrates' dying. But this requires a particular Davidson-type analysis of the identity of actions (1), which may or may not be the true analysis, and it also requires the actual existence of Xanthippe, since the example requires that Xanthippe performs the action of administering the hemlock.

It seems to me that Professor Kim is right in calling Xanthippe's becoming a widow a merely "Cambridge" event, but wrong in thinking that this implies that it is possible seriously to consider the question of whether such an event has any causal relations with any other event or events. It is logically impossible for merely "Cambridge" events either to cause real events or to be caused by them, although they may be one way of describing the logical consequences of real events, given certain relations. There is no need to suppose that ' "Cambridge" dependence' is one way in which an event is non-causally determined by another event or events, or that this shows that the thesis of universal determinism might be true while the thesis of universal causation is false.[1]

University of Liverpool © PAUL HELM 1975

REFERENCES

(1) Donald Davidson, 'Actions, Reasons and Causes', *Journal of Philosophy*, LX, 1963.
(2) Peter Geach, *God and The Soul* (London: Routledge, 1969).
(3) Anthony Kenny, 'Divine Foreknowledge and Human Freedom', in Anthony Kenny (ed.) *Aquinas: A Collection of Critical Essays* (London: Macmillan, 1969).
(4) Jaegwon Kim, 'Noncausal Connections', *Noûs*, VIII, 1 (1974), 41–52.

[1] I am indebted to Raymond Frey for suggestions.

NOTES

The Analysis Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Margaret A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, A. C. MacIntyre, Bernard Mayo, D. H. Mellor, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

Subscriptions. The subscription to Analysis is £2.50 ($7.50) for individuals in Great Britain, £3.00 ($9.00) for individuals abroad, and £3.50 ($10.50) for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford, or placed with any bookseller.

Contributions. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgment is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in Analysis remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

PRINTED IN GREAT BRITAIN BY BURGESS AND SON (ABINGDON) LTD.

ol. 35 No. 5 (New Series No. 167) April 19

ANALYSIS

Edited by
C. J. F. Williams

CONTENTS

Who's at fault? John H. Dreher

Some light on double effect James G. Hanink

Pleasure and the intrinsically desired Jack Nelson *and* David Welker

Remorse, regret and the Socratic paradox C. G. Luckhardt

Jones's body and Marks and Spencer's I. M. Picayune

Reparations to individuals or groups? Alan H. Goldman

Value predications: simple or complex? Samuel A. Richmond

The Kneale-Wiggins regress Ronald E. Nusenoff

BASIL BLACKWELL · ALFRED STREET . OXFORD

ANALYSIS 35.5 APRIL 1975

WHO'S AT FAULT?

By John H. Dreher

SUPPOSE Brown is upset because his friend Smith has just been promoted. Smith has bought a new house and a new car. Brown is so agitated that he goes to his friend Green for help and sympathy. Brown complains that Smith is driving him crazy. As it happens, Green knows Smith too, and Green knows that gentle Smith has always acted with due regard for Brown's interests and feelings. Brown is upset because he is an envious man, and Green says so: he tells Brown that *it is his own fault that he is upset*. Green's remark appears to provide a counter-example to the doctine that a man can be at fault only for what he could have helped, for there is a perfectly good sense in which an envious man cannot help but be upset by his friend's successes.

One reason the case is of interest is that a natural and very common explanation of the putative counter-example is wrong. One might say that what Brown is at fault for is not having corrected his faults. Long ago he should have engaged in a process of reconditioning to rid himself of his envious disposition. Thus we might agree that given that a man is envious, he cannot help but be upset by his friends' successes. Because Brown is envious and so cannot help but be upset by Smith's successes, Brown is not at fault for being upset. Yet Brown is at fault for not having improved his character. But suppose that Brown did try to correct his faults only to find that his envious disposition is ineradicable. Under the circumstances, then, there is nothing Brown could have done to avoid his present troubles. Even so, Green is right: it is Brown's own fault that he is upset by Smith's successes.

It is tempting to say that it is a category error to claim that a person is at fault for being agitated or upset. This intuition is encouraged by the fact that people are generally held responsible only for what they do. Being or even becoming agitated, however, is not something one does; it is not the sort of thing which would show up on one's list of things to do for the day. Becoming agitated is something that happens to a person. Thus, we might conclude that it is senseless to say that Brown is at fault for being upset by Smith's good fortune because being agitated is not something which Brown does.

This line of reasoning is not completely convincing. After all, a person can be at fault for not controlling his emotions (here agitations). It is not clear that controlling one's emotions amounts to much more than concentrating on something other than one's own turmoil, or the cause of it. If one can overcome one's emotions by directing one's attention elsewhere, it would appear that Brown's problem is just that he does not attend to matters other than Smith's successes. Perhaps Brown is at fault for being agitated because he has failed to control his emotions, where controlling one's emotions looks more like something a person does than, say, becoming agitated. On this theory we can say that Brown is at fault for being upset in a kind of derivative sense. He finds no relief from his agitation just because he thinks only about Smith. But this brings us full circle. How can Brown be at fault for being agitated because he failed to control his emotions when, in fact, Brown's problem is just that he is unable to think of anything but Smith's successes?

In any case, we are still left with Green's true remark that it is Brown's own fault that he is upset. That fact seems to conflict with our basic intuition that an envious man cannot help but be upset by his friends' good fortune. How, then, can we allow that Brown is at fault for being upset? I suggest that in 'It is your own fault' we focus our attention on 'own' and not upon 'fault'. What Green is telling Brown in telling him that it is his own fault that he is upset is that *it is not Smith's fault* that Brown's envy has been aroused. Smith did not flaunt his wealth or brag about his promotion. As likely as not, however, Brown holds Smith responsible for his own sorry state. And that is the point of Green's remark. No, Smith was not at fault. But neither is Brown at fault for being upset by Smith's successes, although Brown is upset because he is an envious man. To say it is Brown's *own* fault that he is agitated is not to say that Brown is *at* fault. It is in part to say that Smith is not at fault. It is also to say that Brown's agitation is to be explained by reference to his own faults, in this case his envious disposition. Brown is of course at fault for wrong acts performed out of envy. The point is, however, that it is possible that Brown be agitated because he is an envious man without being at fault for anything.

One might wonder what it is about character defects which makes them defects or faults. What is wrong with being envious? Part of the answer is suggested in the above discussion. There is such a thing as trying to arouse a person's envy. But this does not amount to the attempt to inculcate a disposition; that is, it does not amount to the attempt to make a person who is not envious an envious person. Rather, it is an attempt to elicit certain responses; in fact just those responses typical of a person who is envious. Thus it is sometimes possible to arouse the envy of a person who is not an envious person by artfully contrived acts of provocation.

Because it is prima facie wrong to try to arouse anyone's envy, it is natural for an envious man to think he has been done an injustice even when he has not been provoked at all. After all, it is natural though wrong for an envious person to take his agitation as evidence of wrongful provocation. I suggest that the tendency to mistake innocent behaviour for provocative behaviour is a part of what it is to be an envious person. Certain defects of character are 'extra-punitive'; that is, people who have such faults are too ready to blame others for their own troubles. Faults of pride, such as envy, are paradigms of such defects. They are faults in part because they are dispositions to hold others at fault without just cause.

University of Southern California © JOHN H. DREHER 1975

SOME LIGHT ON DOUBLE EFFECT[1]

By JAMES G. HANINK

SOME philosophers hold (1) that it is always wrong intentionally to kill a human being, whatever the consequences, and (2) that there are absolute human rights, i.e., categorically exceptionless rights. Two such rights may be the passive and negative rights not to be the victim of intentional killing and not to be tortured. I hope these philosophers are correct.

But whatever one's hopes, there are classic conflict cases that seem to mock exceptionless moral principles and absolute rights. Defenders of such principles and rights sometimes try to deal with these hard cases by invoking a principle of double effect. The core of this principle, hereafter the PDE, is fairly clear. One may sometimes blamelessly perform an act having both a good and a bad effect if the latter is merely a foreseen but unintended consequence of one's act. The moral significance of the distinction between intention and foresight is crucial to the PDE.

Still, I am not nearly so clear about the PDE as I would like to be, especially given the philosophical big game I am after. To get at least a bit more clear, I want to examine two recent contributions to the debate over the PDE. They are Mr. Geddes' 'On the Intrinsic Wrongness of Killing Innocent People' (ANALYSIS 33.3, pp. 93–7) and Mr. Duff's 'Intentionally Killing the Innocent' (ANALYSIS 34.1, pp. 16–19). I hope to show that the PDE is not so problematical as they make it, though it is not without its obscurity.

Geddes' contention is that crushing the skull of a foetus to remove it from its mother and so to save her is not an intentional killing. While the foetus' death is certain and certainly foreseen, it is not intended. The obstetrician neither aims for it nor requires it as a means to what is aimed for. Indeed, if somehow the foetus lived, he would rejoice. But if there is in this case no intentional killing, neither is there a violation of the principle never intentionally to kill a human being. Nor is there an infringement of the right not to be the victim of intentional killing. This is crucial. If we suppose that failing a craniotomy both mother and infant will die and that allowing this is folly, it is essential to construe the operation as Geddes does. For if the operation, on the other hand, is taken to be an intentional killing, one must give up the principle and at least one of the absolute rights one might hope to save.

---

[1] My position draws on Germain Grisez's 'Toward a Consistent Natural Law Ethics of Killing', *The American Journal of Jurisprudence* 15 (1970).

Fortunately, I think Geddes' construal of the operation is tenable, though not everything he associates with the PDE is. For he asserts that 'a person may not be held responsible for the unintended evil consequences of his good actions' (p. 95). As it stands this is too strong. Suppose A, a merchant, quite intentionally sells a gun to B. Selling his wares is surely legitimate. But if A foresees B's criminal use of the gun, we would at least sometimes hold A morally responsible for the admittedly unintended evil consequence of his action. If this verdict is correct, it suggests that Geddes' view of the PDE is inadequate or incompletely given. Still, I think his application of the PDE to the craniotomy case is tenable, and I want to look at Duff's argument that it is not.

The argument is a *reductio*. It maintains that if we buy Geddes' application of the PDE, we will be stuck with sophistical solutions to other moral dilemmas. Duff mentions five hard cases. I want to review each to see if we really do get stuck.

First we have the Dudley and Stephens case.[1] Lost at sea on a raft, they eat their cabin boy to avoid starvation. Duff thinks the sophistical solution is to maintain that they might eat the lad 'as a necessary means to the saving of three lives' (p. 17). After all, his foreseen and certain death was neither their end nor a means to it. Surely as it stands this *is* sophistical. But why? Because the PDE, however we finally clarify it, would not allow A and B to kill C and *then* eat him to save themselves. This would not be a case of performing a single legitimate act with both a good and a bad effect. The killing and the eating would be *separate* acts. The latter but not the former could have the good effect aimed at. The killing cannot but be construed as a means to an end and hence inescapably within the agents' intention.

Does this mean that A and B might eat C alive? This would be a single act with both good and bad effects! At the risk of being thought depraved, I suppose there are some scenarios where one would be justified in doing this, if one ate sparingly. C *might* be kept alive until help came. But the Dudley and Stephens case offers no such scenario. For the unique terror the cabin boy would suffer knowing his fate seems great enough to rule out this course of action. There is an injury here, a torturing, quite apart from the killing, that is as much an evil as the killing itself. No comparable terror, I think, would face the crew risking natural death together. Nor, of course, would a foetus experience such terror. If only because of this one might invoke the PDE for a craniotomy but not for Dudley and Stephens.

---

[1] Cf. R. v. Dudley and Stephens; (1884) 14 Q.B.D. 273.

Duff's second case is that of the speluncean trapped in the mouth of a flooding cave and thus blocking the only exit for his equally trapped fellows. The supposedly sophistical solution is to 'blow him free with explosives', with the certain upshot of his death. (Duff offers a more colourful act description.) Of course the death would be no part of the agents' intention. Now one might be uneasy with this solution simply because one wonders why the explorers think the explosive would free the exit and not cover it over entirely. But pressing this isn't playing fair. One might, I suppose, rely once more on the terror factor, but this seems doubtful here, especially if the trapped speluncean's head is outside the cave. Instead I had better swallow the bitter pill. I think the PDE does allow for 'setting the explosive'. Blowing the man free does not, intuitively, seem wrong to me. If the PDE could not be used here, I would be more concerned than I am to find that it can be. I think it just might dissolve the moral dilemma the case presents. It hardly removes the anguish anyone would experience faced with the alternatives. But why should it?

The third case is another medical one. A doctor removes a man's heart and liver and lungs to give three others lifesaving transplants. This is all quite legitimate, says the sophist. The doctor foresees that the first man will certainly die. But his death is not intended nor is *it* required as a means to what is. This "solution" is certainly sophistical. But does it reflect a correct application of the PDE? No. For the PDE hardly allows A to assault B and *then* transplant B's organs to save C, D and E. Again, this is no case of performing a single legitimate act with both a good and a bad effect. The assault itself is illegitimate and, too, it is quite separate from the subsequent transplant operations. So it does not seem that we get stuck with a sophistical solution to this problem case after all.

The fourth case, as Duff handles it, also involves misapplying the PDE. Here a well meaning S.S. man assists in gassing thousands while aiming to save others and get news of the Final Solution to the Allies. This, too, supposedly turns out to be legitimate, since he does not intend but only foresees the thousands of deaths he helps bring about. But again the PDE does not allow A to gas B and a thousand others so that later A, barring the interference of evil men, might save C and even a million others. This is no case of performing a single legitimate act with both a good and a bad effect. The gassing, itself illegitimate, is quite separate from future rescues and revelations, however welcome *their* consequences. Once again it seems that the PDE can be applied much as Geddes does without leading to sophistry in other cases.

The last case, the execution of a scapegoat to stay an enemy attack, is equally unjustifiable. The national leader cannot say he decapitates the scapegoat to satisfy the enemy, thus sparing his people, without intending the scapegoat's death. For the decapitating is one act, the sparing of his people another. Indeed the latter requires the agency of the enemy! He cannot invoke the PDE, for the case at hand is not one in which a single act has both a good and a bad effect. Moreover, there is a strong case to be made that a decapitating just is a killing.

My conclusion, then, is that the *reductio* Duff threatens us with is not so threatening after all. But I can appreciate why he thinks otherwise. Geddes never gives anything like a full account of what the PDE is. Following Anscombe he says that its essence is 'to distinguish between the intended and the merely foreseen consequences of a voluntary action' (p. 94). Quite.

But it is nonetheless only *sometimes* that one may blamelessly do an act having both a good and a bad effect when the latter is a foreseen but not intended consequence of that act. To get any clearer at all about the PDE we must say *when* one can so act. I suggest that at least the following conditions must first be met. Specifying these conditions does not make the PDE crystal clear, but it's a start.

(1) The act must not be wrong anyway, quite apart from its bad effect.

(2) One must intend only its good effect.

(3) One cannot act so that the bad effect is a means to the good.

(4) The act's good effect must be proportionate to its bad effect.

One might invoke the PDE in the speluncean case because all four conditions are met. In the Dudley and Stephens case, (3) is not met if the lad is first killed. If eaten alive (4) is not met. In the transplant, S.S. man and scapegoat cases, (3), if not (1), is not met. It is clear that (3) does a good deal of work for me. Still, it is not the only operative condition. It is (4), for example, that rules out selling a gun for a criminal's use.

Yet mightn't (3) work against Geddes' application of the PDE, too? Isn't the foetus' death a means to saving the mother's life? Well, it certainly isn't a chosen means. Moreover, since its death is only an effect of the doctor's 'wresting it from its mother' it is not itself an action. If one assumes, as I do, that it is first and foremost acts that are the objects of moral assessment, the 'means' to which (3) refers should be understood as itself an act. Evil means are perforce evil acts. Killing the cabin boy, itself a distinct act, would count as a means. So would assaulting one's patient, gassing Jews or decapitating a scapegoat. Thus I do not think (3) is so strong as to rule out the craniotomy; it is strong enough, though, to keep the sophist in line. I hope, to be sure, that both the PDE and especially condition (3) will seem less problematical as philosophers develop more sophisticated theories of act individuation.

Duff is quite right to insist that there are 'logical limits on what I can include in, or leave out of, my descriptions of my intentional actions' (p. 18). One cannot, for example, leave out the means through which one achieves one's aim. This is just why (3) is crucial. But I doubt whether (3), as I see it, would satisfy Duff. Somewhat tentatively he suggests that one logically intends the upshot of an act whenever it is not 'an intelligible *human* possibility' (p. 19) that the upshot *not* follow one's act. The foreseen effect of a craniotomy, the baby's death, seems inescapable. Were it not to happen, we would rejoice. But this is not an intelligible human possibility, though neither is it a logical impossibility.

I have some sympathy for Duff's suggestion. But I think it is suspect. A pair of cases may show why. Suppose A, in the midst of battle, hurls himself on a grenade to save B, C and D. It is not an intelligible human possibility, in Duff's sense, to imagine A's surviving. But I would not say he intentionally or wrongly killed himself. Rather one act, 'hurling himself on a grenade', has a double effect. Only the good effect is intended, although the bad is foreseen. One can justify this deed with the PDE. Consider, on the other hand, the case of A who has his heart, lungs and liver removed to save B, C and D. I am inclined to call this intentional and wrongful suicide. I can account for such very different intuitions about these cases by pointing out that in the second case the operation and the subsequent transplants are separate acts. The former, if construed as a killing, cannot but be seen as a means to a good. Hence it cannot be justified by invoking the PDE, for it fails to meet (3). If someone distinguishes between the operation and its killing effect, indeed a fine distinction, one still can't use the PDE to justify the operation. For that act has itself no good effect proportionate to its bad effect. The good effects await other acts, the subsequent transplants, and even rely on other agents. Thus (4) is not met. The fact that in both the case of the hero and that of the suicide the agent's survival is *not* an intelligible human possibility does not seem crucial to moral assessment.

Michigan State University © JAMES G. HANINK 1975

PLEASURE AND THE INTRINSICALLY DESIRED

By JACK NELSON *and* DAVID WELKER

WE take egoistic hedonism to be the view that the only things intrinsically desired are those things which one perceives as bringing pleasure to oneself. Altruistic hedonism is the view that the only things intrinsically desired are those things one perceives as bringing pleasure either to oneself or to some other person (or creature). It is our view that both kinds of hedonism are false. Pleasure is not the only thing intrinsically desired. We will present several possible cases which, we hope, will show that neither kind of hedonism is a conceptual truth; and if there actually *are* cases of these sorts, as we think there are, then neither thesis is a correct psychological thesis. Consideration of these cases will lead to an analysis of what it is to get pleasure from something.[1]

I

What is it to get pleasure from something? A perusal of the philosophical literature of the past few decades[2] reveals near unanimity on the view that there is no sensation or feeling called 'pleasure', the having of which constitutes getting pleasure from something. That is, even if there are sensations of pleasure, one can get pleasure from something without having such a sensation. We shall not here rehearse the well-known arguments against the view that pleasure is a sensation. One is inclined to feel that its introspective implausibility is the best argument against it.

An adequate theory about what pleasure is should, minimally, take account of the following: (a) We cannot sensibly say things like 'I take great pleasure in sailing, but never have any inclination to sail'; and (b) Often when we say we take pleasure in something or do something for pleasure we mean to draw a contrast with doing it for a further end. Thus we might say 'I garden for pleasure, not for the little extra money it brings in'. While doing something for pleasure and doing it for some further end are neither contraries nor contradictories, each does supply a basis for an *adequate answer* to the question 'Why do you do x?'. These linguistic facts make the kind of analysis of pleasure inspired (at least) by Sidgwick[3] very attractive, viz. 'x gets pleasure out of y' means 'x would rather have y than not, apart from further consequences of having y'.

---

[1] The distinction between what is actually pleasant, what actually brings pleasure, and what is *perceived as pleasant* is obviously crucial. For the sake of stylistic economy we shall frequently ignore this distinction. But we shall not ignore it when we present our analysis in its final form. We shall then use the phrase 'perceive as pleasurable' to cover both cases of presently taking pleasure in something and of regarding some possible state as pleasurable.

[2] See, for example, Gilbert Ryle, *Dilemmas* (London, Cambridge University Press, 1954), Chapter 4, and *The Concept of Mind* (New York, Barnes and Noble, 1948), pp. 107–110; Anthony Kenny, *Action, Emotion and Will* (London, Routledge and Kegan Paul, 1963), Chapter VI; and P. H. Nowell-Smith, *Ethics* (Baltimore, Penguin Books, 1954), pp. 127–140.

[3] See Henry Sidgwick, *The Methods of Ethics* (London, Macmillan and Co., Ltd., 1907),

Now the problem raised by this account of pleasure is all too evident. In effect we have said that to get pleasure from something is to desire it or find it preferable for its own sake. But if we have thereby presented necessary *and* sufficient conditions for getting pleasure from something we have defeated our purpose, for on this account it will be quite impossible to desire anything for its own sake without getting pleasure from it. Yet we want to allow for at least the possibility of someone's intrinsically desiring things other than pleasure. If we are to allow for this possibility, we must regard the foregoing account of pleasure as incomplete. We must suppose that it constitutes at best a necessary, not a sufficient, condition for getting pleasure from something. We need to ask 'What is the difference between getting pleasure from something and desiring something for its own sake while not necessarily finding it pleasant?'

Might one not desire something for its own sake without desiring to *have* it? Considering certain objects of intrinsic desire, it seems not only true, but totally unparadoxical that one should desire that the thing *exist* without desiring that one should possess or enjoy it. Surely I might intrinsically desire racial equality or general moral rectitude; but there is no sense in the idea of my possessing or enjoying such things. On the other hand, it would seem that I could scarcely get pleasure from good wine without desiring that *I* enjoy the wine. And so, as a first shot at distinguishing the pleasant from other things we intrinsically desire, we might suggest that the pleasant is something we desire to have, while the intrinsically desired is something we may merely desire to *exist,* further consequences apart. Obviously, if we desire to have something we also desire for that thing to exist. Consequently, the pleasant is a sub-class of the intrinsically desired, which is a welcome result.

II

When we consider things we intrinsically desire which we *can* possess, must we not *desire to possess* them, rather than merely desiring that they exist? Can I, for example, possibly desire that intellectual consistency and psychological maturity exist without desiring that *I* be intellectually consistent and psychologically mature? We might attempt to allay the doubt raised by this last question by boldly answering the question affirmatively. That is, we might say that obviously I can intrinsically desire that, e.g., intellectual consistency exist without desiring to be intellectually consistent, because I might simply find it too troublesome to be intellectually consistent.

Bk. II, Chapter II. The analysis we propose is in fact a slightly amended version of the formulation found in Alston, 'Pleasure', in *The Encyclopaedia of Philosophy*, ed. Paul Edwards.

As it stands, the above reply to the difficulty is no good because it is equally true that I might find something pleasant and yet still not desire to have it, because it is too troublesome to obtain. Perhaps, though, the distinction we have attempted to draw may be rescued by paying closer attention to the difference between the reasons one might have for failing to pursue something one finds pleasant and the reasons one might have for failing to pursue things one intrinscially desires but does not necessarily find pleasant. Someone might quite conceivably find the prospect of a series of spectacular sensual experiences pleasant without making any attempt to have such experiences. One might, for example, find the prospect of a series of intense sexual encounters and hours of alcoholic euphoria a pleasant prospect. But, made wise and perhaps a bit sad by past experience, our hypothetical character makes no attempt to experience this sensual orgy. His reasons for refraining are the obvious ones: the tarnished reputation, the resultant hangover, the unwanted demands of a personal relationship, and perhaps the prospect of long-term anxiety which seems directed toward nothing in particular. But it is important here that it seems entirely possible that one *could* enjoy the sensual experiences without the unhappy side-effects; which is simply to say that they *are* side-effects. To imagine this, the least we need imagine is progress in the distiller's art, and the most is some not-so-great changes in human psychology. Now if the unhappy side-effects *were* removed from the series of sensual experiences, surely anyone who found them pleasant would have to desire them for himself; he would have to desire to *have* them. He might, of course, still not pursue them because of lack of means, lack of opportunity or the like.

Indeed, we find it beyond our powers to state sufficient conditions for anyone's attempting to enjoy an object of desire. Obviously people do not always do what they want to do or what gives them pleasure. And the factors which can inhibit them from doing so seem so heterogeneous that, for the present at least, we must take refuge in some such catch-all phrase as 'all things being equal'. That is, we may safely say that, all things being equal, people will pursue what gives them pleasure. One, but not the only, thing which can keep things from "being equal" are the unhappy side-effects we have mentioned. Others are a lack of opportunity and a conflict with other desires. In view of these considerations, the contrast we want to draw between the pleasant and the intrinsically desired must be stated in terms of the desires themselves, rather than in terms of what people actually do. We shall provisionally say, then, that if someone finds something pleasant, and its unhappy side-effects are removed, he must desire to possess or enjoy that thing, not merely desire that others have it.

Consider now the contrasting cases in which someone intrinsically desires something but does not find it pleasant, does not pursue it, and does not desire to possess it, even after all its unpleasant side-effects have been removed. Are such cases possible? We want to answer that they *are* possible, on the general grounds that sometimes what one intrinsically desires is such as to have unhappy aspects or consequences which are more than contingently connected with it, are more than *side*-effects. A thing itself may, in certain of its aspects, be repugnant to a person who nonetheless intrinsically desires that the thing exist, that others have it, but not that he have it. To determine whether this is a genuine possibility we need to examine some test cases.

Consider first the matter of psychological maturity. What might this trait involve? Without doubt it would involve indefinitely many more specific personal traits, including consideration for others, possibly concern for others, lack of envy, control of anger, the very minimum tendency to present different "faces" to different people, and a propensity to face the unpleasant rather than indulge in self-deception. As we consider concrete manifestations of these traits, it becomes clear that psychological maturity is very much an ideal rather than anything commonly encountered. It is, for example, evidently quite common to find among academics envy of the salary and reputation of colleagues. Suppose that it were possible to purchase freedom from this envy by coming to believe, whether truly or falsely, that one was somewhat deficient in ability, while still maintaining a healthy self-respect. Of course, it is unlikely that anyone is actually presented with this kind of conscious choice. But if that were possible, it also seems possible that one might choose mild delusion in preference to psychological maturity for the sake of one's own happiness, while desiring that others be psychologically mature, albeit not quite so happy. No doubt a Spinozist would insist that the psychologically mature are necessarily the happiest. But this is far from obvious.

We turn next to moral rectitude. Could one conceivably desire it for its own sake without desiring to *be* morally righteous? To satisfy the requirements of the case we need only postulate a *reluctant* and *regretful* egoist. This is a person who sees that he cannot consistently be both moral and egoistic, and is facing up to the momentous choice between these two kinds of lives. He realizes that the moral life may demand sacrifices of him, and that, having opted for the moral life, if he does not make such sacrifices when required he will be acting immorally. As he carries out a "rehearsal in the imagination" of the moral life he finds that he desires being moral for its own sake; he sees this as the best kind of life. Yet the fact that in its very nature the moral life may require sacrifices overwhelms him. And so he desires that others take up the moral life, but he does not desire that *he* be moral.

Now it may be objected that we have not dealt fairly with this case, on the grounds that our reluctant egoist really *does* desire to be moral, only this desire conflicts with his desire to avoid self-sacrifice. But we think it a mistake to conflate this sort of case with simple conflict of desire; it is really a more complicated sort of situation. Certainly we must, in general, distinguish between having conflicting desires and desiring for others something we do not desire for ourselves. There are clear cases of both sorts to be found. A woman committed to monogamy who must choose between two attractive suitors would experience conflict of desire. But a McGovern supporter who himself had no taste for public life would desire for McGovern and others something he did not desire for himself, namely, public office. Now our reluctant egoist is not *exactly* like *either* of these people. But we suggest that he is *more* like the supporter of McGovern than the woman with two suitors. Our reluctant egoist recognizes that he is so constituted that the moral life is not for him, attractive as it may be. This last phrase 'attractive as it may be' contains the seed of the apparent paradox. If our reluctant egoist finds the moral life attractive, desirable, appealing, etc., how can he fail to desire to be moral? But on the other hand, if opting for the moral life necessarily involves a commitment to possible self-sacrifice, and our egoist finds self-sacrifice abhorrent, how *can* he desire to be moral?

The way out of whatever paradox there is here is simply to recognize that we do not desire for ourselves everything we perceive as desirable. Our reluctant egoist *does* have a *kind* of moral desire for himself; he *desires to desire* to be moral. But, suffering from his abhorrence of self-sacrifice, he does not *desire* to be moral. In effect, he suffers from second-order weakness of will: he fails to have the desires he desires to have; he succeeds only in desiring that others be moral. And in so doing, he satisfies our criterion for intrinsically desiring something without finding it pleasant.[1]

Our final test case is the matter of intellectual consistency. In order to *get* a test case we shall have to assume that knowledge as well as consistency is intrinsically desired, since one obvious way of holding consistent beliefs is to take steps to remain ignorant of any facts inconsistent with a cherished belief. Disavowing any intention of providing an account of rationality, let us christen the desire for knowledge and consistency 'the desire for rationality'. Does, then, rationality in this sense pass our test for being an object of intrinsic desire? Could someone desire rationality for its own sake without finding it pleasant? By now a certain pattern emerges from our discussion which makes this seem entirely possible. We need only suppose that a person cares more for his own happiness than for that of others, or that his happiness alone would be adversely affected by universal rationality, to get the result we want. More concretely, and *very* much more contentiously, let us suppose that someone holds a strong and cherished belief in the existence of God or in the intellectual equality of the races, and that these beliefs are incompatible with admitting the existence of evil and accepting the results of certain psychological studies. Our hypothetical subject then intrinsically desires rationality in the following (required) sense: he desires that others be rational, but he does not himself desire rationality because of the psychological devastation relinquishing a cherished belief would bring upon him. Probably he desires ignorance; but he might conceivably desire to hold two beliefs which he cannot make consistent to himself. Perhaps this latter alternative only makes sense if we suppose our subject holds that his beliefs *are* consistent even though he cannot see how they could be. At any rate, it is easy enough to see how our subject could desire that everyone else be rational. He might know that no one else holds beliefs such that rationality would bring psychological devastation upon them, or else he might not care much about the mental balance of others, so long as they are rational in the sense in question.

---

[1] On second-order desires see Harry Frankfurt's excellent article: 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy*, LXVIII (January 14, 1971).

It would be natural to object here that this case satisfies our criterion for the pleasant as well as that for the intrinsically desired. For if we could but remove the unwanted psychological effects of destroying belief in God or racial equality, would our subject not certainly desire to *be* rational? And could not these effects be removed? But *how* indeed could this come about? There is nothing in the least odd about asking the reluctant abstainer 'Would you not drink if only we could eliminate the hangover?' But there is something decidedly odd in asking the fervent religious believer 'Would you not cease to believe in God, on the grounds that evil exists, if only we could ensure that disbelief in God would not bother you?' No one who intrinsically desires rationality abjures it merely because it conflicts with a belief he has. There is conflict here only when the belief is fervent, edifying, etc. And so the supposition that we could remove the belief without causing psychological trauma is incoherent. Our question to the believer is, in effect 'If I could remove your fervent, edifying belief without removing the fervour and edification, would you not give up the belief in order to be rational?'

III

Having tested our conjecture by examination of the foregoing examples, we conclude that people may intrinsically desire things other than pleasure. In such cases they may desire that these things exist without desiring them for themselves. No doubt such instances will be relatively rare. On the other hand, it is necessarily the case that if someone finds something pleasant and its unpleasant side-effects or consequences are removed, he will also desire to have or enjoy that thing. *Very* roughly, people forgo the pleasant because of prudence; while they forgo the intrinsically desired but non-pleasant because of weakness of will.

We set out to get clearer about what it is to get pleasure from something (to perceive something as pleasurable). So far we have necessary, but not sufficient, conditions for perceiving something as pleasurable. To see that this is all we have so far it will be helpful to put the theses we are defending, as well as the ones we are rejecting, somewhat more formally. In all of the following it is intrinsic desire that is in question.

A: If a person, x, desires something, y, and y is something it is possible for a person to have, then x desires to have y.

B: A person, x finds something, y, pleasant (perceives y as pleasurable), only if x desires to have y (assuming the unfortunate side-effects and consequences of having y, if any, are removed).

We have argued against *A* and for *B*. It is *B* that gives us a necessary condition for someone's perceiving something as pleasurable. To get a sufficient condition for perceiving something as pleasurable it may seem obvious we should accept

C: If a person, x, desires to have something, y, then x perceives y as pleasurable.

But *C* is not an acceptable thesis. Persons who are very conscious of the possible self-sacrifices required by the moral life, but do nonetheless desire the moral life *for themselves* (do not suffer from second order weakness of the will) will satisfy the antecedent condition of *C*. But these people need not find the moral life pleasant.

We might next try

D: If a person, x, desires to have a thing, y, but does not desire to desire to have y, then x perceives y as pleasurable.

We think *D* does give us a sufficient condition for perceiving things as pleasurable, and not just a vacuous one. Sometimes one finds pleasurable things one would just as soon not so find, e.g., embarrassing one's associates in front of their friends or making fun of one's academic and administrative superiors. Such things one will desire to have or do, but not desire to desire to have or do. And these things are correctly judged pleasurable by *D*. Unfortunately we cannot take *D* as the basis for necessary as well as sufficient conditions for perceiving something as pleasurable. Non-ascetics at least normally *do* desire to continue to desire the things they find pleasurable. For example, we find good wine pleasurable and desire both to drink such wine and to continue to desire such wine. Furthermore, there will be cases of overdetermined intrinsic desires. Someone might, for example, desire the moral life for himself both because he perceives it as the best kind of life and because he perceives it as a pleasurable life. (Not all things perceived as pleasurable need be so.) The analysans *D* cannot account for either of these last two kinds of cases. Hence we need minimally a further sufficient condition for perceiving something as pleasurable.

Consider, finally

> *E:* If a person, x, desires to possess a thing, y, and either does not desire to desire to possess y, or *would* desire to possess y even if he did not desire to desire to possess y, then x perceives y as pleasurable.

We *very* tentatively suggest that E just may provide the only sufficient condition we need, and thus also a necessary condition, for finding something pleasurable. Modalities of the type common to A through E are hard enough to interpret; E contains in addition a counter-factual conditional. Roughly, the idea is that if one's desire to possess a thing is independent of any desire to desire to possess that thing one might have, then one does perceive that thing as pleasurable.

Temple University © JACK NELSON *and* DAVID WELKER 1975

REMORSE, REGRET AND THE SOCRATIC PARADOX

By C. G. LUCKHARDT

STATED without qualification the modern versions of the Socratic Paradox are patently false. Claims such as 'Everyone always does what he thinks he ought to',[1] 'A person cannot be said sincerely to think that he ought to do something and not do it, or, again, sincerely to think that he ought not to do something and do it'[2] or 'A man cannot consciously act against or fail to act up to his moral beliefs, judgments or principles'[3] would entail, for example, that a person who presumably thought he ought to attend a faculty meeting, but did not attend owing to a traffic accident on the way, did not really think he ought to attend. Similarly, if these claims are true then a bank president with a gun at his head who opened the bank safe could not really be said to think he ought not to open the safe, and a life-guard could not think he ought to save lives unless he always succeeded in doing so.

On the other hand, a complete rejection of the Socratic Paradox also seems unwarranted, since that would entail that a person could think he ought to attend faculty meetings without ever doing so and with no excusing conditions, and that life-guards could be committed to saving lives without ever attempting to do so. Words could always speak louder than actions.

---

[1] This is R. M. Hare's statement of the Paradox in *The Language of Morals* (Oxford, The Clarendon Press, 1952), p. 169.

[2] P. L. Gardiner, 'On Assenting to a Moral Principle', *Proceedings of the Aristotelian Society*, LV (1954–55), p. 23.

[3] Steven Lukes, 'Moral Weakness', *Philosophical Quarterly*, XV (1965), p. 106.

Clearly the truth must lie somewhere between a wholly unqualified acceptance of the Socratic Paradox and an outright rejection of the claim, in a version qualified by a specific set of provisos. Although various defenders of the claim have acknowledged this, there has been no agreement, nor even a systematic investigation into, what the exact nature of this set of provisos must be. There has, however, been more or less unanimous agreement by both defenders and critics of the claim that any candidate for a true claim will include a proviso excluding such mental states as remorse and regret. Contrary to the evidence that failure to carry out one's beliefs would by itself provide, it is alleged, remorse and regret provide evidence that a person *does* think he ought to act in a way in which he has failed to act. And if this is so, according to many authors, then the existence of such states must be expressly precluded by the Socratic Paradox in order for it to be true.[1]

In this paper I shall argue that, contrary to what is commonly thought, remorse and regret do not constitute evidence against the truth of the Socratic Paradox. Although my strategy in showing this is not to present a fully defensible version of the Paradox, it will involve noting two qualifications that are necessary for the claim's truth, *viz.*, a proviso relating to the time at which the person who fails to act does not think he ought to act, and a proviso relating to the person's ability to act. Once these qualifications are admitted as necessary, it follows that remorse and regret could never occur in any situation to which the claim might be applied, and hence that they are irrelevant to, and unproblematic for, the claim.

As it stands, the claim that

> (A) If a person fails to do what he purportedly thinks he ought to do, then that person cannot really have thought that he ought to do it

is too broad even to be plausible, for it does not state exactly when it is that a person who fails to act in accordance with what he thinks he ought to do does not really think he ought to do that action. It includes too much, for it extends to cases in which one thinks that he ought to do something, but then changes his mind. If this claim were correct, then even in these cases it would follow that the person did not *ever* really think he ought to do what he purportedly thought he ought to do. For this reason the claim must be qualified so as to specify the exact time at which a person could not logically have thought he ought to perform a certain act.

---

[1] Among those who regard remorse and/or regret as relevant to and problematical for the Socratic Paradox are the following writers: R. M. Hare, *op. cit.*, pp. 165–66, 169; P. H. Nowell-Smith, *Ethics* (Baltimore, Penguin Books, 1954), p. 308; H. J. N. Horsburgh, 'The Criteria of Assent to a Moral Rule', *Mind*, LXIII (1954), pp. 345, 354, 355; P. L. Gardiner, *op. cit.*, pp. 28, 38, 41; Irving Thalberg, 'Acting Against One's Better Judgement', in *Weakness of Will*, ed. Geoffrey Mortimore (London, Macmillan, 1971), p. 245; A. Phillips Griffiths, 'Acting With Reason', *Philosophical Quarterly*, VIII (1958), pp. 294–95.

To see what time that could be, imagine the following situation.

> Driving home from work on Friday, Jones notices a large cloth bag in the road. Since cars have been swerving to avoid it, Jones stops to move it out of the way of the traffic. Since it is rather heavy, Jones infers that it must contain something and decides to take it home. If it should contain anything of value and if he can locate its owner, he tells himself, he will notify the owner.
>
> When he arrives home, he inspects the contents and finds that the bag contains over $40,000 in cash. That night he learns that a bank has been robbed of $40,000 and that it is offering a reward of $10,000 for the return of the money. Jones is undecided as to whether to return the money and collect the reward or to keep the money.

Let us suppose that on Saturday Jones decided that he ought to return the money and that he will return the money on Monday. Now if for some reason he should fail to return it on Monday, it does not follow from his failure that he *never* thought he ought to return it. If someone were to insist that this were the case, Jones could reply quite reasonably that he had at one time thought that he ought to return the money, *viz.*, on Saturday, but that he had since changed his mind.

Since Jones' failure to return the money on Monday is perfectly consistent with his believing on Saturday that he ought to return it, it is necessary to limit the time at which failure to act can show that one does not believe that one ought to act. Since Monday's behaviour cannot logically show anything about Saturday's beliefs, and since the same considerations apply to all points of time between Saturday and Monday, a restriction must be put on the probative effect of failure to act. If true, that is, the Socratic Paradox can only be used to show that a person's failure to act is inconsistent with his *present* beliefs about whether he ought to act. Because it cannot be used to show anything about past beliefs, claim (A) should be re-written to preclude such a use:

> (B) If a person fails to do what he purportedly thinks he ought to do, then that person cannot really have thought *at the time he failed to act that he ought to act.*

In addition to this necessary restriction on the application of the Socratic Paradox, however, there is another which is relevant to the supposed evidence that remorse and regret offer.

Suppose that in the example given earlier Jones purportedly thought on Monday that he ought to return the money to the bank, but that on Sunday night a fierce snowstorm had hit the area in which he lived, breaking telephone lines and making it impossible for him to travel or to contact the bank on Monday. Now surely his failure to return the money on Monday would not count as conclusive evidence that he did not think he ought to return it, since Jones could plead that he was *unable* to return it.

Again, suppose that Jones thinks he ought to return the money on Monday, but that when he prepares to leave with the money on Monday morning he discovers that he has misplaced it. At least as long as he is searching for the money, Jones can be said to be thinking that he ought to return it, and yet at the same time he is failing to return it. Here again, his failure to return the money does not count as conclusive evidence that he does not think he ought to return it. Furthermore, even if he were never to find the money, and were never able to return it, this would not show that he did not think he ought to return it.

Or again, imagine a different kind of inability to return the money. Suppose that Jones learns, just as he is preparing to go to the bank, that one of his parents has died. Imagine also that he is so shaken by the news that he finds it impossible to set out for the bank. Here again his inability to do what he thinks he ought to do may disallow the conclusion that he did not really think he ought to return the money.

Whenever a person's failure to act in accordance with what he thinks he ought to do is due to the impossibility of his acting otherwise, claim (B) may not be used to show that he did not think he ought to act. If a person can show that he could not help failing to do what he thought he ought to do, then his claim that he thought he ought to do it may be allowed to stand. Claim (B) can therefore be applied only to those situations in which the person fails to do an action which he purportedly thinks he ought to do and in which he is able to do that action. Although there may be situations in which he is unable to act and is still held not to have thought he ought to act, the reason for holding this will not be his failure, but some other fact. So in order to make it applicable only to those situations in which the person is able to do what he purportedly thinks he ought to do claim (B) must re-written:

> (C) If a person fails to do what he purportedly thinks he ought to do, then that person cannot really have thought at the time he failed to act that he ought to act, *provided that he was able to act.*

We have seen that two relatively uncontroversial provisos must be attached to claim (A) if it is to be true. I have not argued that the claim which results from these additions, claim (C), cannot be false. Hence it is time now to turn to what some philosophers have thought are reasons for thinking that it is false. They have argued that remorse and regret provide evidence that is contrary to the evidence which failure to act provides. They claim, in effect, that although failure to act provides *prima facie* evidence that a person does not think he ought to act, remorse or regret may outweigh that evidence and prove that a person really does think he ought to act. Now while remorse and regret may provide contrary evidence to claims (A) and (B), these claims are, as I have shown, indefensible for other reasons, and must be rejected in favour of (C). Remorse and regret, I shall now show, do *not* provide evidence against (C), and thus are unproblematical for any defensible version of the Socratic Paradox.

In order to see this, it will be helpful first to recall the reason for rejecting claim (A) in favour of (B). (A), it was realized, was too broad, for it enabled one to conclude from a person's failure to perform an action that he *never* thought he ought to perform it. What was required was a revision so that failure to act at a certain time would provide logical evidence only about what a person thought he ought to do *at that time*. This revision was provided by claim (B) and preserved in claim (C).

Given this seemingly necessary restriction on the scope of claim (A), however, it is clear that much of the evidence presented by the existence of such states as remorse and regret is inapplicable in any situation in which claims (B) or (C) might be applied, and hence cannot outweigh the evidence provided by failure to act. Let us consider, for example, the case in which Jones fails to return the money to the bank on Monday, but feels remorse on Tuesday at not having done so. (One cannot of course feel remorse about one's failure to act until after one has failed to act.) In order for such remorse to conflict with the evidence Jones' failure to act on Monday provides, it would have to show that Jones thought on Monday that he ought to act. But this it cannot do. If Jones' remorse on Tuesday could show that he thought on Monday that he ought to return the money, then it would be necessary to wait for remorse *not* to occur before concluding that his failure was inconsistent with his thinking he ought to act. Indeed, if remorse on Tuesday were to show that he really thought he ought to act on Monday, then the conclusion that Jones did not really think he ought to act on Monday could *never* be drawn, since any future state of remorse could presumably override the evidence failure to act on Monday provides. It would always be possible even for Jones himself to suggest that the conclusion that he did not really think he ought to act was premature, since he might in the future feel remorse at his past failure to act.

Since we do not ordinarily think it necessary to wait for remorse not to occur before concluding that a person's professed beliefs are inconsistent with his actions, it follows that at most[1] remorse could only show that a person thought he ought to act at the time he felt remorse and not at a prior time. But while claims (B) and (C) state that a person couldn't have thought he ought to act *at the time he failed to act,* the "contrary" evidence which remorse or regret is sometimes thought to provide would prove that a person thought he ought to act *at a time later than the time he failed to act.* Hence remorse and regret, even if they do tend to show that a person thinks he ought to act, do not provide evidence contrary to the evidence provided by failure to act, and are therefore unproblematical for any defensible version of the Socratic Paradox.

Most such evidence which is alleged to show that despite a person's failure to act, he still might think he ought to act, is of this kind. That is, it is evidence for a person's thinking he ought to do something at a time later than the time of his failure to act. I shall call this sort of evidence 'a *post eventum* criterion'[2] of a person's thinking he ought. The point about the inapplicability of such a criterion to claims (B) and (C) can thus be put in a slightly different way. Not only is the criterion which is alleged to show that a person really does think he ought itself *post eventum,* but a person's thought that he ought, which such a criterion allegedly proves, is itself *post eventum.*

Now although most such evidence which allegedly disproves the truth of claims (B) and (C) is of this *post eventum* kind, not all of it is. Remorse at what one has failed to do must occur after one has failed to do it. One cannot say, 'I feel remorse at what I am doing' or 'I feel remorse at what I am about to do'. Regret too most often occurs after one has failed to act ('I regret that I was unable to attend your party'). In the case of regret, however, it is also possible for it to occur *at the time one fails to act.* One can say 'I regret that I am unable to attend your party'. Likewise one can make excuses for what one is presently failing to do.

Since the argument I have just given for the irrelevance of regret and remorse with regard to the Socratic Paradox has depended on the fact that the evidence which these states present occurs *after* one has failed to act, any evidence presented by regret which occurs *as* one fails to act may still be thought problematical for the Paradox. I shall now argue that even this sort of evidence does not present a problem for a well-formulated version of the Paradox, claim (C).

---

[1] It has been argued that in *no* instance does remorse provide conclusive evidence that a person thinks he ought to act. In 'Remorse' (*Mind,* LXXII (1963), 545–55), Irving Thalberg has argued that since remorse can always be overridden by a contemporaneous failure to act, remorse will not do as an alternative to action, insofar as action is required by one's beliefs. However, this point has been a matter of dispute (see Robert Rosthal's 'Moral Weakness and Remorse', *Mind,* LXXVI (1967), 576–79, and Thalberg's response, 'Rosthal's Notion of Remorse and Irrevocability', *Mind,* LXXVII (1968), 288–89) which I shall avoid here by allowing that remorse *may* tend to show that a person who fails to act actually does think he ought to act. Since on this supposition remorse is irrelevant to the Socratic Paradox, the truth of Thalberg's claim would make remorse irrelevant to the Paradox *a fortiori.*

[2] I have borrowed this phrase from Neil Cooper, who has used it in another context in 'Oughts and Wants', *Proceedings of the Aristotelian Society,* Supp. Vol. XLII (1968), p. 150.

The reason such evidence is not problematical is that when it occurs, there is always something else present which prevents claim (C) from being made, *viz.*, alleged impossibility to act. Suppose that Jones were to claim that at the time he was to return the money to the bank he truly regretted not returning it. Now in order for him to regret it he must, in the first place, actually be failing to return it, since otherwise there would be nothing for him to regret. But in addition to this, Jones must claim, or at least be prepared to claim, that it is impossible for him to return the money. If he were to claim that he regretted not returning the money, and it were pointed out to him that it was quite possible for him to return it, then he could no longer claim that he regretted not returning it. For him to assert that he regrets not returning it when it is fully possible for him to return it is to invite the question 'Then why don't you return it?' In such circumstances there seems to be no answer he can give; the question itself implies that he must not therefore truly regret not returning it.

It follows from this that a person cannot truly regret failing to do an action at the time he fails to do it, provided that he is able to do it. If he is able to do the action, and thinks he ought to do it, then because of the last proviso in claim (C), that claim cannot be applied to any situation in which present regret occurs.

That regret about failure to do a present action can only occur when it is impossible for the person showing the regret to perform the action is reflected in our ordinary ways of expressing regret. Typically, when we express regret at doing (or failing to do) an action in the present, we say that we regret *having* to perform the action (or not being able to perform it). We say such things as 'I regret having to tell you the bad news' or 'I regret that I must inform you that . . .' or 'I regret that it is my duty to . . .' or 'I regret that I am unable to . . .'. Even when we don't express the inability to perform the action explicitly, it is usually implicit in our language. Thus 'I regret to inform you' is equivalent to 'I regret that I must inform you'. Indeed, if the inability is not expressed either explicitly or implicitly, it seems possible always to question the existence of the regret by asking whether impossibility is present. If someone expresses regret at not doing an action, then to ask him why he failed to do it is to ask him to demonstrate his inability to do it. But for this to be the case, regret must imply impossibility.

I have shown that a well-formulated version of the Socratic Paradox can only be defended under conditions of ability. For this reason, regret, even when it occurs at the same time as failure to act, cannot be used to present evidence which would disprove the truth of the claim. Since I have also shown that regret, when it occurs after the failure to act, and remorse, which always occurs then, are not problematical for the claim, it follows that under no conditions are regret or remorse relevant to, or problematic for, the Socratic Paradox.[1]

Georgia State University © C. G. Luckhardt 1975

---

[1] For their helpful comments and advice in the preparation of this paper I am indebted to Professors Robert L. Arrington, Georgia State University, John Beversluis, Butler University, and Nicholas Fotion, Emory University.

JONES'S BODY AND MARKS AND SPENCER'S

By I. M. Picayune

DISCUSSING 'Jones' and 'Jones's body', (or Jones and Jones's body? —actually what was being discussed?—another metaphysical question for another occasion) Bernard Williams says:

> . . . A rather simple argument to the conclusion that they are not the same thing might be attempted on the point that at least 'Jones's body' cannot be substituted for 'Jones' *in the context 'Jones's body'*. But this by itself is obviously a facile objection: an exactly parallel argument could be used to show that 'Marks and Spencer's' does not refer to the same thing as 'the firm of Marks and Spencer's' or 'Marks and Spencer's shop'. ('Are Persons Bodies?' in *The Philosophy of the Body*, ed. by Stuart F. Spicker, p. 147, reprinted in *Problems of the Self*, by Bernard Williams, p. 74.)

I agree with Williams that to appeal to the fact that 'Jones's body' cannot do as a replacement for 'Jones' in the context 'Jones's body' is a rather facile argument for the conclusion that they do not refer to the same thing. But it is nonetheless somewhat complicated to explain what is wrong with it. Williams's reply takes the form of outdoing the argument in facility.

The reason that 'Marks and Spencer's' refers to the same thing as 'Marks and Spencer's shop', is that it, unlike 'Marks and Spencer', has a possessive apostrophe. This key ingredient is lacking in the expression 'Jones', so that the parallel between 'Jones' and 'Jones's body' and 'Marks and Spencer's' and 'Marks and Spencer's shop' is very inexact.

Williams speaks of the phrase 'The firm of Marks and Spencer's', as well as the phrase 'Marks and Spencer's shop', having the same reference as 'Marks and Spencer's'. He probably does not mean that these three phrases have the same reference, but only means that he regards the phrase 'Marks and Spencer's' as being used sometimes to refer to a firm, sometimes to one or other shop belonging to the firm, so that he alleges two distinct counter-examples to the 'facile argument'. A grammatical analysis of the phrase 'the firm of Marks and Spencer's' raises several questions[1] about the rules for indication of possession in English, which, though interesting in themselves, need not be discussed here. It is enough to observe that when this phrase is paired with 'Marks and Spencer's' both members of the pair indicate possession, and so the pair is not analogous to the pair 'Jones's body' and 'Jones'.

This asymmetry regarding possession is a crucial point in favour of the argument as against the spirit of Williams's reply. The pair 'our milkman, Jones' and 'Jones' is another pair where the identity of reference is not disproved by the fact that the first phrase cannot replace the second as it occurs in the first. But in this case, neither member of the pair indicates possession in the way 'Jones's body' does. Grammatically, 'Jones's body' purports to designate something possessed by the thing designated by 'Jones'. A proper reply ought to argue either that this 'possession' is a (grammatical?) illusion or that in some cases the possessor may be identical with the thing possessed. Williams's examples do not apply to either of these projects.

---

[1] I am indebted to my colleagues W. K. Goosens, John Marshall, James Cargile, Daniel Devereux and Cora Diamond for discussions which made me aware of the complexity of these questions.

University of Virginia © I. M. PICAYUNE 1975

REPARATIONS TO INDIVIDUALS OR GROUPS?

By ALAN H. GOLDMAN

JAMES Nickel bases his latest argument (ANALYSIS 34.5, pp. 154–60) for reparations to groups upon a distinction between the justifying and administrative basis for a programme of reparations. I have argued against compensatory hiring for groups elsewhere,[1] but had not considered there this most recent argument. Its novelty lies in the shift from abstract or ideal principles of compensatory justice to the necessity in practice of balancing claims so as to maximize (imperfect) justice. The justification for favoured treatment for groups according to Nickel derives from the administrative feasibility of such a programme by comparison with the high cost and impracticality of administering compensatory justice in this area on an individual basis. Thus while there is only a high correlation between being black, for example, and having been discriminated against and so deserving compensation (justifying basis), so that preferential treatment for the group will occasionally result in undeserved benefits for individuals, the balance of justice in practice favours such treatment. The viable alternatives seem to be either award of deserved compensation in the great majority of cases and occasional undeserved benefit and hence injustice to white job applicants, or compensation on an individual basis, which would require demonstration of past injustice in court or before a special administrative body, so that the cost and difficulty of the operation would result in far fewer awards of deserved reparation. It is better, the argument holds, to have compensation which is only almost always deserved than a programme which in practice would amount to almost no compensation at all, so that a policy which would not be accepted in an ideally just world (a world which became ideally just after compensation was paid) becomes best in the present situation.

---

[1] 'The Justification of Reverse Discrimination', presented at the Eastern Division meeting of the American Philosophical Association, December 27, 1974.

In reply, one would first of all like to ask how high the correlation between group membership and past discrimination, and hence the proportion of deserved to undeserved compensation, must be. Presumably a ratio of fifty-one to forty-nine will not do, since in the case of compensation by preferential hiring policies, there are two injustices involved in every case of undeserved compensation: first the payment for the undeserved benefit made by society in accepting less efficient service (since the candidate will not be as competent if hired only because of preferred treatment), but, more important, the injustice to the white male applicant who is best qualified.[1] The correlation is presumably not as high in the case of women as a group as in the case of blacks, and not as high for middle class as for lower class blacks. The latter comparison suggests narrowing the specification of the group so as to maximize the correlation, but of course at the limit of such narrowing is a programme administered on an individual basis.

Thus far it still seems we must balance ideal theory against practice, but the far more serious point completely forgotten in Nickel's argument is the effect of the operation of market criteria upon hiring even within a compensatory programme of preferential treatment. Since hiring within the preferred group still depends upon relative qualifications and hence upon past opportunities for acquiring qualifications, there is in fact an inverse ratio established between past discrimination and present benefits, so that those who benefit most from the programme, those who actually obtain jobs, are those who deserve to least. Given that those individuals will always be hired first who have suffered least from prior discrimination, this effect of competence requirements completely destroys the rationale of arguing by correlation unless the correlation is extremely close to perfect, for as long as there are some members in the market who have not unfairly lost opportunities, they will be the ones getting the jobs. But the establishment of such high correlation for a specific group, or the narrowing of specifications for group membership until virtually all members have been treated unjustly, amounts to administering a programme on an individual basis. It will have to be determined for each individual whether he belongs to the narrowly specified group and have to be determined for individuals within that group whether virtually all have been discriminated against and thereby suffered harm.[1] These two steps, when the group is sufficiently narrowly specified, will I suspect be as difficult as handling cases on an individual basis from the beginning.

---

[1] Philip Silvestri has used an argument (ANALYSIS, 34.1, p. 31) similar to Nickel's to justify only voluntary reparations. But since preferential hiring is never voluntary for those white males who apply for and do not get the jobs, his position is irrelevant to that crucial issue.

Since Nickel's argument wrongly assumes that the majority of compensation cases will tend to be fair if the correlation of past injustice to group membership is above fifty per cent, it is unsound. Nor can the practice of hiring by competence itself be blamed or held therefore unjust. If efficiency is Nickel's basis for arguing for group compensation, he cannot condemn a general practice which in the long run results in more goods for all in favour of some less efficient alternative. It would at least be strange to recommend hiring the least competent as a general practice within some preferred group, and in this case efficiency would be gained not lost by moving away from a group programme. In order then not to create a policy which in practice singles out for benefits within a generally unjustly treated minority just that minority which has not been unjustly treated, and thus does treat unfairly members of the 'majority group' applying for jobs, a compensatory programme of preferential hiring must be administered on an individual basis. Where there are significant departures from this toward preferential treatment for loosely defined large groups, we must suspect further injustice not only to 'majority group' members, but to members of the minority who have suffered previous injustice and are now passed over in favour of other members who have at least suffered less. There are surely degrees of injustice, and the market here will invert the ratio of past injustice to present compensation if the programme is directed toward a group.

University of Idaho © Alan H. Goldman 1975

---

[1] I suspect that for the past generation of blacks this might be found to be the case, but not for the generation currently in school, and certainly not for women, who are benefiting most from the current practice in universities.

VALUE PREDICATIONS: SIMPLE OR COMPLEX?

By SAMUEL A. RICHMOND

IN arguing for the simplicity of 'good' Moore presented us with a trichotomy: good is either complex, simple or meaningless. His argument then proceeded by eliminating the first and third alternatives. I shall try to show that his argument fails because he failed to eliminate the first alternative. This failure was due to a misconception about complexity. He did not consider the possibility of *value complexity.* Moore tried to prove the absolute simplicity of 'good' through arguments that it was unanalysable. I shall try to show that the supposition that predications of 'good' are value complex accounts for Moore's arguments as well as the hypothesis that they are absolutely simple. I shall argue that if there are no value simples, that is, if all value items are value complex, then 'good' is unanalysable in non-value terms. It follows from this that it is possible that good is unanalysable and complex. In this case Moore's assumption that if good is unanalysable it must be simple is false.

A predication is absolutely simple if its truth or falsity is independent of every other predication. Moore believed 'yellow' was absolutely simple. A value predication is value simple if its predication is independent of every other value predication. Moore believed intrinsic value simple. A predication is complex if its truth or falsity is dependent upon another predication. Moore believed 'horse' was complex. A predication is value complex if it is dependent upon another value predication. Moore believed 'right' was value complex; he was a teleologist.

Given the preceding, a predication might be value simple and still be absolutely complex. It might be independent of any other value predication and be dependent upon other non-value predications. Moore would view naturalist definitions of intrinsic value as taking this view. It should be also noted, however, that if a value predication is not value simple, then it is not absolutely simple.

We can view naturalism as an attempt to analyse all value predications in terms of value simples. That something is conducive to the survival of the race is dependent upon other non-value predications but not upon other value predications, according to naturalists who maintain 'good' means *conducive to the survival of the race.* Thus, they believe there is a fundamental, value simple but absolutely complex sense of 'good'. Moore argues that 'good' is not analysable into other value terms or non-value terms. He insists upon the absolute simplicity of goodness—that each predication of goodness is independent of every other. (The goodness of a whole is not a function of its parts, and vice versa.) There can be no question that he believed in its simplicity as simplicity has been defined here. He wrote:

> And so it is with all objects, not previously known, which we are able to define: they are all complex; all composed of parts, which may themselves, in the first instance, be capable of similar definition, but which must in the end be reducible to simplest parts, which can no longer be defined. But yellow and good, we say, are not complex: they are notions of that simple kind, out of which definitions are composed and with which the power of further defining ceases. (*Principia Ethica,* Cambridge University Press, 1968, pp. 7–8.)

Suppose there are no value simples. Suppose for every value predication *V* there is another value predication *W* such that *V* is dependent upon *W*. Suppose, for example, that for every state of affairs the value of that state of affairs is dependent upon the value of its consequences and that there is no state of affairs that fails to have consequences of value. In this case there would be no value simples, and as a consequence all naturalistic analyses would be impossible. But the impossibility of analysing values in terms of things value simple but absolutely complex, would not imply the absolute simplicity of value. Quite the contrary, the unanalysability into non-value terms would be due to the value complexity of values. And, of course, the predications in question would be absolutely complex because value complex.

This shows that Moore's assumption that if good is not analysable in terms of non-value terms that are absolutely complex but value simple or analysable into other predicates both absolutely and value simple, then good is absolutely simple, is false. His arguments do not establish the absolute simplicity of good. For the failure of the analysis into value simples may be due to the non-existence of value simples. All value may be value complex.

Are all value predications complex? In the preceding it has been argued that if they are all complex then none are analysable in terms of value simples. But we have not established whether they are complex. I shall not attempt to do so in this paper. I would like, however, to indicate things that lend the hypothesis sufficient plausibility to make it worthy of investigation. First, it has already been indicated that certain statements of teleology, or consequentialism, might already have built into them a commitment to the complexity of value predication—those that view the value of a state of affairs, *S*, as a function of the value of the states of affairs which are the consequences of *S*, for every *S*. Second, universalization theses are also committed to the complexity of value predications in singular judgements. For if one believes that no state of affairs *S* is good unless all other states of affairs of the same type are good, then the truth of every predication of good in a singular judgement is dependent upon the truth of an indefinite number of other such predications.

Moore's concern with simplicity and complexity was not an idiosyncrasy. In the *Tractatus,* Wittgenstein was much concerned with these ideas. He argued that for terms to be meaningful they must be fully determinate in their meaning; and to be fully determinate, they must be analysable into absolute simples (2.02–2.0212). Moore seems to anticipate this view in assuming that if 'good' cannot be analysed into simple terms, 'good' itself must be simple. Such a theory of meaningfulness would not permit it to be the case that all value predications were both value complex and meaningful. This may explain why Moore concluded basic value terms were absolutely simple when he found they could not be analysed into non-value terms. It may also explain in part why logical positivists who could not accept value simples among absolute simples rejected value terms as not cognitively meaningful.

Quine's argument regarding the indeterminacy of meaning is in the exact opposite direction from Wittgenstein's argument from meaningfulness to simples (cf. *Word and Object*, chapter ii, especially section 9). Quine argues that since there are no simple predications, none the truth or falsity of which is independent of all others, there is an indeterminacy of meaning. No predication is related to reality independently of all other predications. As sentences come closer and closer to being conditioned solely to non-verbal stimuli they approximate to simple predication. But Quine believes simplicity is a matter of degree and that no sentence is absolutely simple. He believes that in general we cannot isolate the stimulus conditions of a given sentence from the stimulus conditions of that sentence taken in conjunction with intrusive information. For any given sentence S with stimulus conditions C there will be another sentence S' with stimulus conditions C' such that S comes to be conditioned to S' and to C'. Quine believes that it is impossible given the combined sets C and C' to say which part constitutes the stimulus meaning of sentences relevant to the truth of S but not part of the stimulus meaning of S. Furthermore no matter how firmly conditioned a given sentence S may be to conditions C, its conditioning to other sentences and via them to other conditions may yield a weakening of the conditioning of S to C when all the other conditions elicit responses inconsistent with S.

Wittgenstein claimed determinacy of meaning implies simplicity of predication; Quine claims complexity of predication implies indeterminacy of meaning. The latter is the contrapositive of the former; they agree upon the basic relationship between the simplicity or complexity of predication and the determinacy or indeterminacy of meaning. They disagree on what direction the argument should take. Should we establish simplicity by requiring determinacy of meaning, or prove indeterminacy by recognizing the complexity? The meaningfulness of value terms turns on the difference. Wittgenstein's early view taken with Moore's arguments against naturalistic analyses provides support for viewing value terms as absolute simples (intuitionism) or for viewing them as meaningless (noncognitivism). Quine's more recent view of the matter opens the

My purpose here is not to heap further criticism upon Frege's early treatment of identity. On the contrary, I wish to defend it against what is perhaps the only objection which does not work. This objection, as first put forward by William Kneale, questions the possibility of interpreting identity as a relation between signs.

> For if we say that the phrase 'the morning star' has the same content as the phrase 'the evening star', is not this equivalent to an assertion that the content of the phrase 'the morning star' is identical with the content of the phrase 'the evening star'? And if it is, must we not go on to say that our original statement is not about the phrases 'the morning star' and 'the evening star' but about the phrases 'the content of the phrase "the morning star" ' and 'the content of the phrase "the evening star" '? But even this is not enough; for the proposed solution involves an infinite regress. (Kneale, *The Development of Logic,* Oxford, 1962, p. 494.)

The answer to both of Kneale's questions is in the negative, for his argument turns on an equivocation on the reading of the identity sign. He correctly notes that in the *Begriffsschrift* ' "≡" was to be understood as a symbol for *identity of content* between expressions' (ibid.). This leaves his argument in the form of a hypothetical equation between (1) the morning star ≡ the evening star and (2) the content of 'the morning star' ≡ the content of 'the evening star'. If we accept (2) as being equivalent to (1), then the regress is started. However, to do this, we must read the '≡' in (2) as 'is identical with', whereas the '≡' in (1), as Kneale points out, is read as 'is identical in content with'. Furthermore, keeping the regress going turns on reading '≡' differently in any one sentence from one step to the next, e.g., to get from (2) to the next step, the '≡' in (2) would be read, not as it is read when going from (1) to (2), but rather, as it was then read in (1).

Kneale's confusion in setting forth this objection in the first place seems to be shared by David Wiggins, who also finds an infinite regress. Fortunately, it is easier to pick out this confusion in Wiggins' formulation, for rather than letting confusion slide into equivocation by equating (1), (2), and so on, *ad infinitum*, Wiggins sees each step of the regress as an attempt at reaching an explanation of the previous one.

> Asking for the sense of 'a = b' I am told 'a' and 'b' have the same content, or designate only one thing. Unless something is said to justify calling a halt here, the explanation generates a new statement of the same form as the original explicandum . . . 'The content or designatum of "*a*" = the content or designatum of "*b*".' Applying the same explanation to this we get . . . [same procedure as Kneale, above]. (Wiggins, 'Identity-Statements', in R. J. Butler (ed.), *Analytical Philosophy: Second Series* (Oxford, 1965), p. 51.)

That something need be said 'to justify calling a halt' is quite puzzling. The only reason Wiggins gives for starting the process in the first place seems to be that the original explicandum is not only about the signs themselves. 'But evidently we never can reach in this way [the procedure given above] what seems to be needed to carry the explanation through, a statement *only* about signs' (ibid.). Why should we require a statement 'only about signs'? Rather, what is needed is a statement asserting a relation between signs. Perhaps Wiggins' (and Kneale's) point is that the relation of sameness between the contents themselves is left unexplained. However, despite what is said by both Wiggins and Kneale, on Frege's view of identity as a relation between signs nothing is said about any such relation (between contents), for it is only the signs themselves, and not their common content, which are being related.

Frege's criticism of his early view was that it does not allow identity sentences to be used to express any 'proper knowledge' about objects. What such identity sentences do express is knowledge about the signs as designators, and about objects only as designated by the signs. The common content of the signs (i.e., the object) is not asserted to be self-identical (rather, this is the extent of the 'proper knowledge' that his later view is generally taken to allow identity sentences to express), for in fact it is this common content through which the signs are related. Still, there may be a bit more expressed about an object by an identity sentence interpreted as relating two signs.

For Frege, a sentence is a complex name which expresses a thought (its sense), and this in turn presents a truth value (its reference). The reference of a sentence is a function of the reference of its parts, and if one of its parts has no reference, neither does the sentence as a whole (cf., e.g., 'On Sense and Reference', pp. 62–3). For an identity sentence to have either the True or the False as its reference (or more simply, for the identity sentence to be either true or false), each of its parts must have a reference. For such an identity sentence to be true, the content of both signs must be one and the same object. Though on Frege's early view of identity the signs stand for (or designate) themselves in an identity sentence, it must still be the case that the signs themselves have a content other than themselves. Otherwise, they could not be used to form *any* sentence having reference.

Therefore, to assert an identity sentence is to assert that there is some, i.e., at least one, object such that it is the content of both '*a*' and '*b*'.[1] Looked at in this way, there is no inclination whatsoever to require either a statement 'only about signs' or an explanation of any relation between the contents of the signs. Hence the regress never even gets started.

The Ohio State University © Ronald E. Nusenoff 1975

---

[1] One cannot object here that the inclusion of the assertion of existence renders all identity sentences contingent. The interpretation of identity as a relation between signs has already done that!

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Margaret A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, A. C. MacIntyre, Bernard Mayo, D. H. Mellor, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

SUBSCRIPTIONS. The subscription to ANALYSIS is £2.50 ($7.50) for individuals in Great Britain, £3.00 ($9.00) for individuals abroad, and £3.50 ($10.50) for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford, or placed with any bookseller.

CONTRIBUTIONS. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgment is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in ANALYSIS remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

26 DEC 1975

PRINTED IN GREAT BRITAIN BY BURGESS & SON (ABINGDON) LTD., ABINGDON, OXFORDSHIRE

35 No. 6 (New Series No. 168) June 197

ANALYSIS

Edited by
C. J. F. Williams

CONTENTS

A note on the essence of natural kinds — Colin McGinn
Resolving a paradox of inductive probability — Kenneth S. Friedman
On entailment — P. T. Geach
What *de re* belief is not — Michael Corrado
Rorty's new mark of the mental — R. I. Sikora
Swinburne on confirmability — R. I. Sikora
Are there art-critical concepts? — Joel Rudinow *and* R. I. Sikora
Determinism and knowledge — William Lyons
Hart on paternalism — Christine Pierce
Locke's unnatural kinds — Douglas Odegard

ASIL BLACKWELL . ALFRED STREET . OXFORD

ANALYSIS 35.6 JUNE 1975

A NOTE ON THE ESSENCE OF NATURAL KINDS

By Colin McGinn

I WISH to take up an issue arising out of J. L. Mackie's piece 'Locke's Anticipation of Kripke' (ANALYSIS 34.6). It concerns the question which properties (if any) of a natural kind—more especially a substance, such as gold—are to be reckoned *essential* to the kind; i.e., which properties actually instantiated by a kind are such that *that* kind could not have lacked them, or, in another idiom, which properties a given kind instantiates in all worlds in which it exists.

After remarking Locke's observation that we are prone to (try to) annex substance terms, e.g. 'gold', to the *real* essence of the substance in question, rather than as epistemological modesty requires to its *nominal* essence, Mackie writes:

> It is a consequence of our using the word 'gold' with this intention that if we contemplate the counterfactual possibility that something with this same internal constitution [i.e. real essence] was not (through some change in other things or in the laws of nature) shining yellow in colour, malleable, fusible, soluble in *aqua regia*, and so on, we would express this by saying that gold might not be yellow, might not be malleable, etc., whereas if we contemplate the counterfactual possibility that something with a different internal constitution had all these features, we would say not that gold might have a different internal constitution, but only that something else might look and behave like gold. Saul Kripke has argued just this, that if gold does in fact have atomic number 79, it is not possible that *gold* should have a different atomic number, though it is possible that gold should lack the features by which we now recognize it, and that some other substance should have them (pp. 178–9).

Now turning to Kripke, we read:

> In particular, then, present scientific theory is such that it is part of the nature of gold as we have it to be an element with atomic number 79. (We may also in the same way, then, investigate further how colour and metallic properties follow from what we have found the substance gold to be: *to the extent that such properties follow from the atomic structure of gold, they are necessary properties of it* [my italics], even though they unquestionably are not part of the *meaning* of 'gold' and were not known with a priori certainty.) (Pp. 320–1, 'Naming and Necessity' in D. Davidson and G. Harman, eds., *Semantics of Natural Language* (Dordrecht: Reidel, 1972).)

It seems plain enough from this passage that Kripke is asserting the negation of a thesis defended by Mackie and attributed by him to Kripke. For Mackie says (and says that Kripke says) that, on the one hand, the atomic structure, as indicated in the atomic number, of (e.g.) gold is to be reckoned to that substance's essence and, on the other, that the various dispositional (and other) properties of gold are merely contingent; while Kripke claims that the latter properties, provided that they 'follow from' the atomic structure, are to be reckoned essential too. I want to ask who is right and why. Before tackling the issue directly, though, I pause to consider two prior matters, which may serve to clear the air for the main issue.

First, a comment on Mackie's account of the provenance of those 'necessities of constitution' he seems willing to acknowledge. His suggestion, in a nutshell, is that their provenance is to be located in a certain utility attaching to the 'linguistic policy' noted in connexion with Locke, i.e. the policy of annexing substance-terms to the property or properties comprising the real essence of the substance and prescinding from those comprising its nominal essence. That Mackie regards the adoption of such a policy as up for decision is borne out by his remark that we *could* have chosen to fix the reference of 'gold' via its nominal essence, as Locke recommended. Then, according to Mackie's account, the properties of gold which are now supposed accidental to it—malleability, fusibility, etc.—would be counted essential. Furthermore, the property in which that substance's real essence consists—sc. having atomic number 79—would on this alternative policy be counted contingent. So the modal properties of a substance are consequent upon our selection, prompted by considerations of utility, of a subset of the substance's properties, which thereafter fix necessary and sufficient conditions for something's being that substance. Since we could have selected a different subset, the possibilities concerning the substance could have been different; in particular, we could have counted the colour, etc., of gold essential and the atomic number inessential. (Reducing a putative necessity to an agreed contingency, namely our linguistic policies, is bound to entail that there is after all no necessity about the necessity.) Actually Mackie's account seems to have a still less palatable consequence. For suppose we decided, for whatever reason, to annex 'gold' to such properties of gold as that of being a stuff endowed by human beings with symbolic significance, or being a medium of exchange, or being what the Queen's crown is made of: then it would follow, from Mackie's principles, that these properties would be reckoned to gold's essence, while its other properties, whether pertaining to real or nominal essence, would come out contingent. That is, it would be supposed necessary and sufficient for a substance to be gold that these properties were satisfied by it. But now, since rubber could have been the stuff instantiating *these* properties, it seems to follow that gold could have been rubber, and contrariwise. Or at least that we would say, and say correctly, that gold might be (or might have been) what we now call rubber; in other words, on the alternative policy, we could truly say, *of* gold and rubber, that the one might have been the other, since the other might have instantiated the properties to which we originally annexed the name of the one. I infer that the prospects are not good for founding the modal properties of natural kinds on (alleged) features of our 'linguistic policies'. The boot is surely on the other foot.

It may seem surprising (and this is the second matter) that Kripke should make the parenthetical remark quoted above. For, it may be protested, hasn't he spent painstaking pages showing that the identifying marks by which we initially fix the reference of a natural kind term, e.g. 'gold', are neither necessary nor sufficient for something to be of that kind: i.e., that it is *possible* for something to be gold despite its lacking those marks? On closer inspection, however, the apparent tension is easily resolved. For the possibilities *here* envisaged by Kripke, and deployed to advance a certain thesis about the sense of terms like 'gold', are to be interpreted *epistemically,* in Kripke's sense. Hence his talk of optical illusions and other sources of error. Indeed, Kripke observes that, in *this* sense of 'possibly', gold might have turned out (or might turn out) not to have atomic number 79. But in the passage cited above the possibilities under discussion are *metaphysical,* again in Kripke's sense. So the diagnosis of Mackie's misinterpretation, as of the apparent tension, is that epistemic modalities are being taken as metaphysical; once separated out, there is no inconsistency. What could have turned out to be the case compatibly with the reference of 'gold' staying fixed as *that stuff* and what could have been, given what is, are just different questions. And the quoted passage speaks to the latter, while the epistemic possibilities discussed bear on the former. (In fact, this is made very clear at p. 319 of 'Naming and Necessity', and I hesitate to labour the point.)

Now to the main issue. I begin by sharpening up the model of natural kinds, in particular, substances, that Kripke and Mackie (and before them Locke) are working with. The leading idea of the model is that properties of a kind come in two grades, one more fundamental than the other. Let's call properties of the more fundamental grade *primary* nomic properties, and let's call properties of the less fundamental grade *secondary* nomic properties. Then the point of the model is this: that the primary properties—e.g., having a certain atomic number—are 'responsible for' and 'underlie' the secondary properties—e.g., solubility, conductivity, etc. As it were, the secondary traits of a substance are nomically derivative from, because they have their basis in, the primary traits. Explanation of a substance's behaviour as manifested in its secondary properties, consists in discovering its internal *nature*; and its nature—what the stuff fundamentally *is*—is constituted by its primary properties. We know what the nature of a substance is when we have identified its primary properties and seen how they account for the more readily discernible secondary properties.[1] Now, in terms of this model, we can state the issue more amenably: Are the secondary properties of a substance to be reckoned to its essence along with the primary properties, or are they not? Is the relation between properties of these levels contingent or noncontingent? If the relation is noncontingent—or rigid, as I prefer to say—then Kripke is right; if it is contingent—or nonrigid—then Mackie is right.[2] I think we can adduce considerations favouring Kripke's answer to the question thus stated.

Mackie might insist that a condition sufficient for the relation's being *non*-rigid is satisfied. (See his '*De* what *Re* is *De Re* Modality?', *J. Phil.*, Vol. LXXI, No. 16, p. 552.) For, on his view, the relation between the atomic structural properties of gold and its superficial properties is *causal*; and, since causal relations are contingent, so is the relation between the two sets of properties. It is contingent because the identity of the items causally related at a time doesn't turn on their being so related at that time. Now I am inclined to agree that *if* the relation were causal in this way—i.e. a relation holding between distinct items—then it would have to be nonrigid. But I deny that the relation we are concerned with is causal in this way, for I deny that primary and secondary properties are distinct in the required way.

I think we can arrive at a correct view of the relation by combining two thoughts, both of which enjoy a certain currency. The first is, as Kripke notes, that scientific inquiry proceeds by *reductive identification*—of water to H_2O, of heat to molecular motion, and so on. The second is that this procedure takes a specific form: it consists in finding a basis in the infrastructure of a substance for the superficial properties characteristic of the substance. Bringing these two thoughts together, we reach the thesis that investigation of the nature of a substance involves *identifying* the secondary properties with the primary ones. More especially, the dispositional properties of a substance are explained by identifying them with underlying 'scientifically fundamental' properties, where these latter may be characterized dispositionally or non-dispositionally, depending on predilections not here in question. So for gold to have the property of being soluble in *aqua regia just is* for it to have some atomic structural property. Basic among these properties will be, as physics has discovered, the atomic number of gold. Uncovering the primary properties of gold is finding out *what it is* for gold to have the secondary properties it nomically has. And this is obviously a relation more intimate than causality, as ordinarly construed.

---

[1] Cf. David Wiggins' concept of 'scientifically fundamental properties', which occurs in 'Essentialism, Continuity, and Identity', in *Synthese*, Vol. 28, Nos. 3/4, November 1974. Note that the distinction of grades here sketched is not to be identified with the classical distinction between primary and secondary qualities, at least on some construals of that distinction. It has more in common with the Aristotelian distinction between properties (in the old, narrow, sense, translating '*idia*') and essence.

[2] I picked up the term 'rigid relation' from M. K. Davies; it is intended to recall Kripke's concept of a rigid designator, according to which the relation between a name and its bearer is a rigid relation—if it holds in the actual world, it holds in all worlds.

Now if this view of the relation is correct, we are a short step from the rigidity we set out to establish. For, as Kripke and others have made clear, true identities are necessary: accordingly, property identities, though discovered a posteriori, are necessary. If secondary properties are to be *identified* with primary properties, then they will be necessarily coextensive—i.e., they must be related rigidly. Whence it follows that if gold couldn't have lacked its atomic properties, it couldn't have lacked its dispositional properties either. For to lack the one would just *be* to lack the other. That, I suggest, is the reason Kripke's laconic remark is true. The nomic *equivalence* of primary and secondary properties precludes any peeling off of dispositional properties; for to peel off the dispositional properties would be to peel off the atomic characteristics too. (Note that this thesis isn't committed to a reduction of the *sense* of predicates ascribing secondary properties to that of predicates ascribing primary ones; we have to do with different 'modes of presentation' of the same property.)

There is a further consequence of this view. It is that the symmetry of identity ensures that rigidity cuts both ways. If the atomic constitution cannot come apart from the superficial traits, then neither can the superficial traits come apart from the internal constitution. This consequence may seem to contradict Kripke's remarks about real gold and fool's gold. For isn't this precisely a case in which two substances agree in superficial properties, yet diverge with respect to molecular constitution? Well, gold and iron pyrites agree in *some* secondary properties, and to that degree they share some primary properties. But Kripke nowhere claims that they agree in *all* non-molecular properties; and of course they do not. The situation rather seems to be this: if substances S_1 and S_2 agree in secondary properties $\phi_1, \ldots \phi_n$, then there are primary properties $\psi_1, \ldots \psi_n$ in which they also, and necessarily, agree on account of property identity. So, if S_1 and S_2 coincide on *all* secondary properties, then they coincide on all primary ones. In particular they would perforce agree in atomic structure. Otherwise put, if the properties of S_1 and S_2 coincide, as physically described (i.e., in a certain vocabulary), then they coincide under all other descriptions. For example, if the properties of S_1 and S_2 fall under the same chemical descriptions, they would fall under the same physical descriptions. We are dealing with identical states of a substance, variously characterized. My claim is that *if* this view is right, and it seems to me that it is, then Kripke's position on the modality of secondary properties is correct and Mackie's incorrect.

Nor is the position unintuitive. Indeed it enables us to undercut what may have seemed embarrassing questions. On the Mackie view, which seems common, at least among sympathetic readers of Kripke, the following has to be swallowed. Consider a possible world containing a substance with all the secondary properties of gold yet lacking its atomic constitution, and suppose also that this world contains a substance having that constitution yet lacking all of gold's secondary properties, suppose indeed that it instantiates all of the secondary properties of rubber: then it is the second substance not the first that is gold. On the view I defend, however, such implausibilities are avoided, since on that view the world just described is impossible. It *isn't* possible for something to instantiate the molecular properties of gold and instantiate the dispositional properties of rubber, and vice versa. For to instantiate the molecular properties of gold *is* to instantiate its dispositional properties. The situation is perhaps clearer in respect of "natural phenomena" such as heat and light. We may simply reject the possibility that heat (= molecular motion) or light (= stream of photons) should exist in a world yet not heat or illuminate things, respectively; indeed that they should, for example, have the secondary properties of water or oxygen and behave as *they* do. On the nonrigid view of the relation between basic nature and observable properties, however, such possibilities must, it seems, be countenanced; and the identity of the substances in question is to be fixed simply by the underlying properties, so that, e.g., light could have had the dispositional properties of water. According to the view I defend, on the other hand, we are free to regard gold, heat, light and water, in all possible worlds, just as we always did—i.e., as endowed precisely with those properties that intuition supposed definitive of the kind. (This is not to deny that the scientifically fundamental properties enjoy a certain primacy among properties of a kind. It *is* to deny that this status amounts to unique essentiality.)

Someone may object at this point and say 'But now isn't it going to follow that *all* of gold's properties are essential to it, and isn't this false?' I don't believe my principles do have this profligate consequence, however. For first, not all of gold's properties are *nomic* properties of it, e.g. being a medium of exchange, or being deposited in Fort Knox: i.e., it is not a consequence of the laws governing gold that it have these properties. And second, such nomic properties of gold (or more clearly heat and light) as that it produces certain phenomenologically identified sensations in human beings don't come out essential either (granted certain assumptions about the modal properties of human beings) for there is no question of *identifying* such sensations with the primary properties of gold (or heat and light). The kind and the sensation really are distinct existences, and so are contingently related. Rigidity sets in just when the properties in question are identifiable, i.e. when the predicates ascribing them represent identical states of the substance.

A final word about Locke. A not untypical passage from the *Essay* is this:

> I doubt not but if we could discover the figure, size, texture, and motion of the minute constituent parts of any two bodies, we should know without trial several of their operations one upon another; *as we do now the properties of a square or triangle* (my italics, Book Four, III, 24).

Perilous though it is to read current preoccupations into past philosophers, I conjecture that, if anything, Locke would have concurred with the rigid view of the relation between real essence and 'powers', as manifested in the substance's relational properties. (Note also his remark that 'malleableness depends on, and is *inseparable from*, the real essence of gold', quoted by Mackie at p. 177.) 'Scientifical' knowledge, were it available to us, would be, though a posteriori, of necessary truths. That, in modern dress, is the claim I have tried to ground.

University College London © COLIN McGINN 1975

RESOLVING A PARADOX OF INDUCTIVE PROBABILITY

By KENNETH S. FRIEDMAN

SUPPOSE (1) $P(a|b_1) = P(a|b_2) = \ldots = P(a|b_n)$. Then the reasonable odds at which to bet on *a*, given that either b_1 or b_2 or . . . or b_n is true, are unaffected by knowing that in fact it is b_1 that is true. Since at least one of these statements is true by hypothesis, and since it does not matter which one is true, one can arbitrarily assume the truth of any one. This argument is attractive, and Stephen Spielman ('Assuming, Ascertaining, and Inductive Probability', in *American Philosphical Quarterly Monograph #3: Studies in the Philosophy of Science* (1969), pp. 143–161) uses it with effect to criticize the plausible and widely held rule (C) that the probability of a claim based on the total available evidence relevant to that claim determines the reasonable odds at which a person should bet on that claim.

Spielman considers two fair dice being rolled, the claim, *a*, that the total rolled is 7, and the sextuple of claims b_n, that at least one n (n ranging from 1 to 6) is rolled. Since $P(a|b_1) = P(a|b_2) = \ldots = P(a|b_6) = 2/11$, condition (1) is satisfied. So, argues Spielman, reasonable odds at which to bet on *a*, given that any of the b_n is true, are 9 : 2. However, $P(a|b_1 \vee b_2 \vee \ldots \vee b_6) = P(a) = 1/6$, so reasonable odds on *a*, according to rule (C), are 5 : 1.

For the purposes of this paper I will consider two fair coins being flipped. The only difference between this example and Spielman's is that this is a simpler one, considering only two, instead of six, possibilities. Let *a* be the claim that the total number of heads is one. Let b_1 be the claim that at least one coin lands heads and b_2 be the claim that at least one coin lands tails. Now $P(a|b_1) = P(a|b_2) = 2/3$, again satisfying condition (1). So according to Spielman's argument reasonable odds at which to bet that exactly one coin lands heads, given that either at least one coin lands heads or at least one coin lands tails, are 1 : 2. However, since $P(a|b_1 \vee b_2) = P(a) = 1/2$, the reasonable odds according to rule (C) are 1 : 1.

In these examples, in which the b_n are mutually exhaustive, Spielman's argument can be strengthened. For any n (n = 1, 2), if b_n then the probability of *a* is 2/3. However, if there exists an n (n = 1, 2) such that b_n, then the probability of *a* is 1/2. These two previous statements, both obtained from rule (C), are mutually inconsistent, which implies the inconsistency of probability theory. (In Spielman's example n ranges from 1 through 6, and we obtain the following mutually inconsistent pair: For any n, if at least one n occurs then the probability of obtaining a 7 is 2/11; If there is an n such that at least one n occurs, then the probability of obtaining a 7 is 1/6.)

Spielman concludes that rule (C) is defective, and that the inductive logics, such as Carnap's, for which it serves as a basis are similarly defective. I wish to defend rule (C) by arguing that Spielman's argument, though attractive, is not valid.

The example may help bring out the problem in Spielman's argument. Note that b_1 and b_2 are mutually exhaustive, so that we know that b_1-*or*-b_2 must be true. Now suppose we learn that in fact b_1 is true. This does constitute new information, since $P(b_1|b_1 \vee b_2) = 3/4$. But why should it change our estimate of the likelihood of *a*? Simply because the only state of affairs which b_1 rules out (two tails) is one in which *a* is false. Similarly, if we had learned that b_2 is true, this would change our estimate of the likelihood of *a* because the state of affairs ruled out by b_2—though not the same as that ruled out by b_1—is also one in which *a* is false. So either knowing that b_1 is true or knowing that b_2 is true rules out only a state of affairs in which *a* is false. However, knowing that either b_1 or b_2 is true is not the same, and it rules out nothing.

The point is that even though we know that either b_1 or b_2 is true, and even though it makes no difference which is true, we cannot arbitrarily assume one of them to be true. The reason for this is not that we might assume that the wrong one is true. It is rather that assuming that a particular one is true increases the probability (or reasonable odds) that *both* of them are true. On the assumption that b_1 or b_2 is true, $P(b_1 \& b_2) = P(b_1)+P(b_2)-1$. On the assumption that b_1 is true, $P(b_1 \& b_2) =$

$P(b_2)$. Thus the assumption that b_1 is true increases the reasonable odds (in the sense of increasing the likelihood) that both b_1 and b_2 are true, provided b_1 is not certain ($P(b_1) < 1$); and cannot decrease the reasonable odds that both b_1 and b_2 are true. Yet the truth of both b_1 and b_2 can radically affect the reasonable odds on *a*. Thus it is neither surprising nor a defect of rule (C) that that rule should yield the result that $P(a|b_1) \neq P(a|b_1 \vee b_2)$ in the above example. For the conditioning information is different in the two cases; moreover, it is different in a way which affects the probability of b_1-*and*-b_2, which, in turn, affects the probability of *a*. So Spielman's argument is invalid. (Spielman presents other examples than the one I have considered, but these can be handled along the same lines.)

Note, too, that the statement 'If b_1 then the probability of *a* is 2/3' is not a correct translation of $P(a|b_1) = 2/3$. For that putative translation entails 'If both b_1 and b_2 then the probability of *a* is 2/3' which is false. So in the case in which we know that at least one heads has occurred or we know that at least one tails has occurred a better translation would be: 'For any n, if b_n and there is no relevant information other than b_n then the probability of *a* is 2/3.' Similarly, a better translation for the case in which we know that at least one heads occurs or at least one tails occurs is: 'If there is an n such that b_n and there is no relevant information other than there being an n such that b_n, then the probability of *a* is 1/2.' These two statements are mutually consistent, since in the former case there is additional relevant information beyond the fact that there is an n such that b_n. So it is not shown that rule (C) is inconsistent.

Note that if the b_n are mutually exclusive then this criticism of Spielman's argument is inapplicable. However, in that case his argument does not lead to results incompatible with rule (C).

State University of New York
at Oswego

© KENNETH S. FRIEDMAN 1975

ON ENTAILMENT

By P. T. GEACH

WRITING about the theory of entailment I have put forward, Briskman[1] supposes me to have been trying 'to capture the intuitive notion of *B*'s logically following from *A*', and accordingly infers from my theory 'the . . . result that although we can validly *prove* q from p & $\bar{p}$, it is not the case that q logically *follows* from p & $\bar{p}$'. (His emphases.)

I first wish to make a terminological distinction between a proof and a derivation. In my way of speaking, if you offer a *proof* you are asserting the premisses and the conclusion; but you can *derive* conclusions from unasserted premisses, merely assumed for the sake of argument. The notion of what logically follows applies to derivations generally, not only to proofs. I am not sure how clear Briskman is about this matter: he not only speaks of using 'p & $\bar{p}$' as a premiss in a "proof", he actually speaks of "asserting" this premiss! Even if Lewis and Langford did the like, this way of speaking is still objectionable.

My view about valid derivations was and is that their validity depends on the possibility of exhibiting them as series of steps such that the premisses and conclusion of each step are related by a relation that I called entailment: the name does not matter, but my account of the relation was near enough to the accounts given by others who have used the same word for my use of the word not to be seriously misleading. On my account it turns out that the relation is not unrestrictedly transitive. There is no need to worry about this: a logical chain is sound iff every link is sound, and my criterion of entailment determines the soundness of each link, but there is no reason why such a link should *directly* join the first premiss to the final conclusion.

Briskman of course has not shown that when a Geachian entailment holds there is a conclusion not logically following from a premiss, nor yet that there is a conclusion that follows from a premiss but cannot be joined to it by a chain of Geachian entailments. Still less has he produced formal objections to my formal account (*Logic Matters,* pp. 186–188). I conclude that his objections are merely rhetorical.

Part of Briskman's paper illustrates, I fear, the aptness of Quine's remark that the plainest writing is not proof against stalwart reading. I have several times insisted on the unrestricted transitiveness of valid proof; I have insisted on this even with a certain brusqueness in dismissing the opposed view (*Logic Matters,* pp. 63, 184, 211). In face of this, Briskman devotes several paragraphs to the supposition that perhaps I did not mean what I said; needless to say, I did mean just that, and I have not changed my mind.

---

[1] 'Classical semantics and entailment' by Larry Briskman in ANALYSIS 35.4.

My talk about relations in the present note is not wholly appropriate, in my view, as may be seen from my first paper (see *Logic Matters,* pp. 175–177), I regard 'entails', not as a relative term joining names, but as a propositional connective. This doctrine is well developed, and the grounds for it set out, in Arthur Prior's *Objects of Thought* (pp. 26–29, 53–56 f.).

The derivation of 'q' from '$p \,\&\, {\sim}p$' with which Kielkopf[1] seeks to embarrass me is of course not a valid derivation in plain propositional calculus, *PC*, since some of its steps are wffs whose main connective is an arrow for 'entails', and no such wffs occur in *PC*. Construing Kielkopf's work a little more kindly, one might take the entailment theorems he cites at steps (2), (3), and (9), not as premisses, but as annotations, showing that the three steps from (1) to (4), from (1) to (5), and from (8) to (10) are steps corresponding to Geachian entailments. But in that case Kielkopf's construction simply reduces to one of Lewis's, and calls for no further comment.

Reconsideration of the article by Dale that Briskman cites has suggested to me that a small change in my criterion for entailment may be desirable. The original criterion would still serve for deciding whether there is an entailment-theorem in the extension ES of a formal system S corresponding to a material-implication theorem in S, *so long as the antecedent of the theorem is not disjunctive in form nor yet its consequent conjunctive.* If we did get an implication-theorem in S; with a disjunctive antecedent or conjunctive consequent, it would have to be split up into simpler implications. Thus suppose $\ulcorner CAA\alpha\beta\gamma K\rho\sigma\urcorner$ is a theorem of S; for a normal system S, $\ulcorner C\alpha\rho\urcorner$, $\ulcorner C\alpha\sigma\urcorner$, $\ulcorner C\beta\rho\urcorner$, $\ulcorner C\beta\sigma\urcorner$, and $\ulcorner C\gamma\rho\urcorner$, $\ulcorner C\gamma\sigma\urcorner$, will all be theorems of S; and by the new test I am proposing, it will be a theorem of ES that $\ulcorner AA\alpha\beta\gamma\urcorner$ entails $\ulcorner K\rho\sigma\urcorner$—assuming that none of α, β, γ, is again a disjunction, and neither ρ nor σ is a conjunction—just in case, by the *old* test, $\ulcorner\alpha$ entails $\rho\urcorner$, $\ulcorner\alpha$ entails $\sigma\urcorner$, $\ulcorner\beta$ entails $\rho\urcorner$, $\ulcorner\beta$ entails $\sigma\urcorner$, $\ulcorner\gamma$ entails $\rho\urcorner$, $\ulcorner\gamma$ entails $\sigma\urcorner$, are all of them theorems of ES. This modification removes certain "paradoxes", without seriously complicating the test procedure or the proof that if theoremhood in S is decidable so is theoremhood in ES; I need hardly say that it makes no difference to my general view of the way entailments are involved in valid arguments.

University of Leeds

© P. T. Geach 1975

---

[1] 'Adjunction and paradoxical derivations' by Charles F. Kielkopf in ANALYSIS 35.4.

WHAT *DE RE* BELIEF IS NOT

By Michael Corrado

CONSIDER the sentence

(A) The man who is sitting is believed by John to be standing.

(A) is meant to be understood *de re,* if anything is. To understand it *de dicto,* we would have to suppose that John has a contradictory belief, namely the belief that the man who is sitting is standing. However, there is at least one other sense in which it can be taken: the man who happens to be sitting is believed by John to be standing. Clearly there is a difference, then, between the *de re* and the *de dicto* readings of the sentence. No one, I believe, would dispute that. The question that is of interest is whether all *de re* belief propositions are really capable of being rendered solely in terms of *de dicto* belief propositions. A number of analyses have been proposed that claim to show that such a rendering is possible;[1] but it seems to me that they all turn out to be analyses of another notion, which I will call accidental reference. In this paper I want to explain what accidentally referential belief propositions are and to argue that accidentally referential propositions are not *de re.*

Suppose, to begin with, that both

(B) John believes that the one who is standing must pay the bill

and

(C) The man who is standing is Governor of New York

are true. Suppose further that (B) is only true *de dicto.* Perhaps John's belief amounts to this: that there is exactly one man who is standing, and whoever is standing must, by the rules, pay the bill. Does anything follow, now, from the truth of (B) and (C)? We know that names for the same thing need not be substitutible one for the other in a sentence expressing a *de dicto* belief proposition, *salva veritate.* For example, we should not be able to substitute 'the Governor of New York' for 'the one who is standing' in (B). Nevertheless, it seems to me that the truth of

(D) John believes that the Governor of New York must pay the bill

does follow from the truth of (B) and (C). The trick is to understand in what *sense* (D) must be true. It is important to see that if I, knowing (B) and (C) to be true, assert the truth of (D) to a friend, I am attempting to convey considerably less information than could be conveyed if (D) were straightforwardly *de dicto*, and, furthermore, I am not likely to be misunderstood. Of course if my friend takes me to be saying something that *is* straightforwardly *de dicto*, he *will* misunderstand me; (D) understood *de dicto* is probably false, given the truth of (A). Hence (D) understood *de dicto* does not follow from (B) and (C). If he understands it in the sense in which I intend it, on the other hand, he will not misunderstand me; and in that sense the truth of (D) does follow from the truth of (B) and (C).

---

[1] See, for example, Wilfrid Sellars' 'Some Problems about Belief', in J. W. Davis, *et al.,* eds., *Philosophical Logic* (Dordrecht: D. Reidel, 1969), pp. 46–65. But the analysis I have in mind especially is one formulated by Roderick Chisholm in a reply to a paper of mine, both forthcoming in a volume on Chisholm's philosophy.

But what is this sense in which (D) must be true if (B) and (C) are? Perhaps I have just been asked who must pay the bill; I obtain from John the information that he believes that the man who is standing must pay the bill, and I know that the man who is standing is the Governor. What I convey by means of (D) is this: if John is right in his belief, then it is the Governor that must pay the bill. What is John's belief, precisely? I haven't said. How is his belief related to the Governor, precisely? I haven't said that either. What I have said, in fact, comes to something like the following, in which the belief is to be understood *de dicto*

> (E) For some description of the Governor and for some proposition expressed by a sentence which asserts that the individual fitting that description must pay the bill, John believes that proposition.

If we do understand (D) in some such way as this, (D) does indeed follow from (B) and (C). And in fact we do sometimes understand sentences like (D) in this way. I would say that 'the Governor of New York' is *accidentally referential* in (D); and I would say about propositions that can be rendered by sentences like (E) that they are accidentally referential belief propositions. Obviously (B) would turn out to be accidentally referential, in a trivial way.

To return to our question, then, are all *de re* propositions merely cases of accidental reference? And are all accidentally referential propositions *de re*? Is the proposition expressed by (D) *de re*, for example? Might it be better expressed according to our conventions by means of the sentence

> (F) The Governor is believed by John to be the one who must pay the bill?

If there is some difference between what is expressed by sentences like (D) and what is expressed by sentences like (A) and (F), then we should be able to say what it is. Even if we succeed in showing that there is some difference, of course, we will not have shown that the *de re* is not somehow reducible to the *de dicto*. We will have shown at most that some analyses which have purported to be analyses of the *de re* have not been.

What I want to do first is to show that *de re knowledge* propositions are not merely accidentally referential; and then to argue from the relationship that knowledge bears to belief that the same is true of belief. We would ordinarily say that if the Governor is known by John to be the one who must pay the bill, then John knows who must pay the bill; that is, we would consider it inconsistent to say that the Governor is known by John to be the one who must pay the bill, and yet John doesn't know who must pay the bill. But any *de dicto* proposition, including the proposition that John believes that the Governor must pay the bill, is consistent with John's not knowing who is to pay, however odd it may seem. We can imagine John saying to himself, 'Now if only I can find out who the Governor is, I will know who is to pay the tab'. In the extreme example, that the man who must pay is known by John to be the man who must pay is not consistent with John's not knowing who must pay; but that John knows that the man who must pay is the one who must pay is consistent with his not knowing who must pay. And in general John may know that the F is the G without knowing who the G is. Here are two more examples, which we can imagine have been excerpted from a conversation with a police investigator

> (G) We know that the occupant of this room is the killer. Now we have to find out who occupies this room.
>
> (H) All we know at this time is that the cheque forger is Clarence Bush; we've yet to find out who Clarence Bush is.

In both cases the first sentence expresses knowledge *de dicto*, the second expresses knowledge *de re*.

Even if we knew an extended list of *de dicto* propositions to be true, like the following

S knows that the Y is the Z
S knows that the X is the Y
S knows that the W is the X
Etc.

we could not be sure that S knows who the Z is, or that he knows anyone to be the Z. What is required is that at some point we introduce a *de re* assertion: that the W is known by S to be the X, for example. But every one of the sentences in the list expresses an accidentally referential proposition, in a sense which is easily derived from the notion of accidentally referential belief propositions: the proposition that S knows that the A is F is accidentally referential iff it is to be understood thus

> For some description of the A, for some proposition expressed by a sentence which asserts that the individual fitting that description has the property F, S knows that proposition to be true.[2]

And if all of the sentences in the list are accidentally referential, and none of them *de re,* it follows that for knowledge at least the *de re* and the accidentally referential are two different things.

An example might be more helpful here. If I overhear the conversation from which (G) was excerpted, and I know that the occupant of this room is your brother, I am justified in telling you that the police now know that your brother did it; I am justified, but only if what I tell you is accidentally referential. But because I am justified in telling you that in the weak accidentally referential sense, it does not follow that the *de re* proposition expressed by

> (I) Your brother is known by the police to be the killer

is also true. If it were, we would have to agree that the police know who the killer is; and by their own admission they don't. Hence in this case the *de re* is something different from accidental reference.

Unfortunately for our discussion of belief, there is no such thing as *believing-who*. But we can say this, I think: if the Governor is believed by John to be the one who must pay, and if the Governor *is* the one who must pay, and if there is the proper relationship between what John believes and the evidence, then the Governor is known by John to be the one who must pay. And since *de re* knowledge propositions must be distinguished from knowledge propositions which are merely accidentally referential, it follows that *de re* belief propositions must be similarly distinguishable from belief propositions which are merely accidentally referential. There is no amount of evidence that will get us from

> (D) John believes that the Governor of New York must pay the bill

to

> (J) The Governor is known by John to be the one who must pay the bill,

since all of the evidence that could get us from (D) to something that John knows must be evidence for what John believes in (D), and although we have no way of telling it from (D) itself, what John believes is that the one who is standing must pay the bill. For we have understood (D) to be accidentally referential, derived from (B) and (C). We might be able to get from (D) to (J) with the help of

---

[2] There are some problems with this formulation and the previous one, but I think it is clear enough for my purposes. In particular I want to avoid the device of substitutionally-interpreted quantifiers. Besides descriptions we might equally well have made use of some sort of individual concept.

(K) The Governor is known by John to be the one who is standing;

but as it happens *no one* is known by John to be standing; he believes of the man who is sitting that he is standing. But if (D) were *de re* in addition to being accidentally referential, then, as we have already seen, since the Governor *is* the one who is standing, the proper sort of evidence, properly placed, would get us from (D) to (J). So in the case of belief, as well as in the case of knowledge, the *de re* is something different from accidental reference.

It may seem that I have done a great deal of writing to establish what is perfectly obvious, that the proposition expressed by (D) (understood as I have indicated) is different from the proposition expressed by (F). But I have already said that analyses that promise to explain the *de re* in terms of the *de dicto* often turn out to be analyses of sentences like (D), and thus of accidental reference. Though I am not prepared to argue here that the *de re* is irreducible to the *de dicto*, I have tried to contribute something to the discussion by marking off the *de re* from something that it is not.

Ohio University © MICHAEL CORRADO 1975

RORTY'S NEW MARK OF THE MENTAL

By R. I. SIKORA

PROFESSOR Rorty has claimed ('Incorrigibility as the Mark of the Mental', *The Journal of Philosophy,* June 25, 1970) that the only thing that distinguishes events in the stream of consciousness (henceforth 'mental events') from physical ones is that sincere first-person reports of mental events are treated as incorrigible; he argues that because of this, mental events may eventually "disappear" even though life somehow continues much as it is now. I have argued ('Rorty's Mark of the Mental and His Disappearance Theory', *Canadian Journal of Philosophy,* September 1974) that incorrigibility is not the mark of the mental; and if it is not, the view that it is obviously gives no support to his disappearance thesis. In response to this, Rorty has offered a new mark of the mental ('More on Incorrigibility', ibid.). I shall argue that his new theory of what distinguishes the mental is incompatible with his view that mental events may eventually disappear.

My argument against the claim that incorrigibility is the mark of the mental is that after-images are mental yet sincere first-person reports about them are *not* treated as incorrigible. If a person reported that his after-image had 12 sides and the figure responsible for it had in fact 13, it would surely be in order to challenge his report.

Rorty grants my point, and offers a new mark of the mental. A statement that *p* asserts the occurrence of a mental event iff 'there are no accepted procedures for resolving doubt [whether] *p*, given that [the report that] *p* fits into a pattern of sincere reports made by a subject S, even though it fits into no more general theory'. The key expression is 'pattern of sincere reports'. On his old theory, something would fail to qualify as mental if there were procedures for challenging a single report about it (as there are for after-images). On his new theory, after-images can qualify as mental even if there are procedures for challenging a single report as long as there are no procedures for challenging a multiplicity of reports. And Rorty claims, I believe correctly, that there are no such procedures. He suggests that if a person repeated his claim about an after-image a number of times, we would probably give in and regard the case as grounds for doubting our general theory correlating the number of sides of an after-image with the number of sides of the physical object responsible for it.

It would seem then that Rorty has provided a new mark of the mental that will withstand my counterexample. But he has avoided one problem only to create another. He wants a mark of the mental that will enable him to claim that at some time in the future mental events may cease to have the characteristic in question and so will in a sense disappear, and that this could happen without any important change in our lives. This view, the disappearance thesis, will seem more plausible if we remember that his mark of the mental is what may be called an extrinsic characteristic. Thus the events in question could continue to have the same intrinsic characteristics (as would be required for there to be no important change in our lives) yet cease to be mental provided that they lost the extrinsic characteristic that we have been considering. Rorty defends this view in the following argument (he has generously agreed in a letter that the argument I sketch is the only argument he has to offer):

> For something to be paradigmatically mental, it must be private. For something to be private, it must be such that one and only one person can make incorrigible reports about it. At some time in the distant future, scientists may make such great advances in the study of the brain that there will be grounds for challenging reports about the stream of consciousness. Then challenges to first person reports of mental events will be in order, and such reports will cease to be incorrigible. As a result of this, the stream

> of consciousness will cease to be private and therefore will cease to be mental. Mental events will have "disappeared".

The trouble is that with his new mark of the mental the argument no longer works. He formerly held that at some time in the future, scientific advances may justify us in overruling a sincere first-person report about the things we now call sensations; and that at that time they will cease to be sensations because first-person reports about them will cease to be treated as incorrigible. With his new mark of the mental, the situation is quite different. For intense pains, for example, to cease to be mental we would have to be prepared on the basis of a scientific theory to reject not merely single but repeated first-person reports of intense pain. We would have to respect the verdict of a scientist observing something in a person's brain—perhaps the firing of a certain C-fibre—*even if* the scientist's verdict conflicted with repeated claims on the part of a sincere person that he was undergoing intense pain. But it seems clear that the grounds for rejecting the scientist's claim would be at least as good as the grounds Rorty now has for rejecting the claim of a scientist that a person was repeatedly miscounting the sides of his after-image. Thus, replacing '*p*' with 'I am experiencing intense pain', there would continue to be 'no accepted procedure for resolving doubt about the existence of an intense pain given that its existence fits into no more general theory'. So intense pains would still bear the mark of the mental; they would not have "disappeared".[1]

This is true despite the fact that if the theory correlating pain with brain states were correct, there would probably be no occasions when sincere reports of an individual regarding intense pain conflicted with the claims of a scientist observing his brain. This is the case because what counts for Rorty is not whether such conflicts would actually occur but whether there would be a procedure for resolving them in favour of the scientist if they did, and it seems obvious that there would not.

---

[1] If the person only claimed that he was in mild pain or claimed that he had a pain that was only *slightly* different from the pain that the scientist attributed to him, we might agree with the scientist; in this case, Rorty would be in a position to claim that at least some sensations had disappeared. But if the existence of mental events is incompatible with materialism, a materialist obviously needs to get rid of *all* of them, not just some of them.

The University of British Columbia © R. I. Sikora 1975

SWINBURNE ON CONFIRMABILITY

By R. I. Sikora

PROFESSOR R. G. Swinburne argued, in a note in this journal, that the confirmationist principle is false because a certain sort of statement that is obviously meaningful is nevertheless neither confirmable nor disconfirmable ('Confirmability and Factual Meaningfulness', Analysis, 33.3, pp. 71–6). I criticized his counterexamples to confirmationism on the grounds that the sort of statement he considered was in fact capable of being confirmed or disconfirmed ('Confirmability and Meaningfulness', Analysis, 34.4, pp. 142–4). In response to this criticism, Swinburne has offered a new counterexample to confirmationism ('Meaningfulness without Confirmability—a Reply', Analysis, 35.1, pp. 22–7). He suggests (p. 24) that any statement of the form p is unconfirmable:

> p: Among possible claims about the prehuman past to which the best evidence ever to be obtained by man gives probability x some are true.

He allows us to substitute for 'x' any probability we like.

Unfortunately, for I share Swinburne's desire to show that a statement can be meaningful *without* being confirmable, statements of form p seem to be confirmable. The assertion is that some claims about the prehuman past that will have a definite given probability when the best evidence is in are true. But the more claims there are which have the degree of probability in question when the best evidence is in, the more probable it will be that at least one of them is true. And it would seem that we could get evidence in regard to how many claims are likely to meet this probability requirement. At any rate Swinburne would have to show that we can't conceivably get such evidence in order to substantiate his claim that p is unconfirmable.

My point may be illustrated by a comparison with a game of chance that is like a lottery in all respects save that there is no guarantee that anyone will win. In the sort of game in question, although there is no guarantee that anyone will win, there is a guarantee that each person who bets will have a certain *chance* of winning. Now without showing that any given person will have a different chance of winning from that guaranteed, we can get evidence in regard to how likely it is that one person at least will win by getting evidence as to how many persons will buy tickets for the game. Our ticket holders are meant to be counterparts of statements having a certain probability of being true when the best evidence is in; winning is meant to be a counterpart of being true.

The University of British Columbia

© R. I. Sikora 1975

ARE THERE ART-CRITICAL CONCEPTS?

By JOEL RUDINOW *and* RICHARD I. SIKORA

WE wish to discuss a position whose influence on theory of criticism has been considerable, despite a good measure of fairly obvious implausibility. The position, baldly put, is this: It is impossible to apply an item of art-critical terminology truly and without change in meaning to two distinct aesthetic objects; or put more baldly still: It is impossible to say truly the same art-critical thing about two distinct aesthetic objects. Since a concept, by its nature, must necessarily be multiply instantiable, the position amounts to a denial that art-critical terms correspond to art-critical concepts. We therefore refer to the position as the non-conceptual or NC thesis for short. Further, if we say that to describe something is to bring it under a universal, or to say that it instantiates a concept, it follows that the features one attributes to an object in describing it must be shareable ones. The NC thesis then comes to this: In art criticism one does not, strictly speaking, describe the work of art; those of its features which are the particular concern of the critic (let us say its aesthetic features) are not shareable.

It will be convenient for us to discuss the NC thesis with reference to a widely read article by Arnold Isenberg, entitled 'Critical Communication',[1] with which the NC thesis is not uncommonly identified. We too shall treat Isenberg as therein committed to the NC thesis, but further as so committed by way of a *non-sequitur*. Though we think that there is sufficient textual evidence for our reading of Isenberg on this point, the fairness of this reading is not absolutely crucial. Isenberg in any case does not, as we shall, indicate a way of avoiding the NC thesis while still doing justice to those apparent facts about art criticism which make the NC thesis attractive.

A number of writers have hinted at the non-conceptual character of art criticism, often in connection with the idea that works of art are in a special way unique and unrepeatable, or that it is in the nature of a work of art to be—not "one of a kind"—but not of a kind. There is a family of peculiarities about the practice of art criticism which appears to motivate such speculations.

First, there is the proliferation of figurative language in art criticism, from which, perhaps, it is to be learned that in criticism insinuation, rather than direct description, is the order of the day. For example, Lucy Lippard writes of Cleas Oldenburg (using, we take it, some of his words)

---

[1] In *Aesthetics and the Theory of Criticism: Selected Essays of Arnold Isenberg* (Chicago: University of Chicago Press, 1973), pp. 156–171.

> Painting, which has slept so long in its gold crypts, in its glass graves, is asked to go out for a swim, is given a cigarette, a bottle of beer, its hair rumpled, is given a shove and tripped, is taught to laugh, is given clothes of all kinds, goes for a ride on a bike, finds a girl in a cab and feels her up . . . (*Pop Art,* New York, Praeger, 1966, p. 108).

A less self-indulgent, but equally typical example comes from Harold Rosenberg's *The Anxious Object* (New York, New American Library, 1966).

> [I]n many of the landscape-figure-abstractions of the last few years [de Kooning] has achieved through speed a lightning clarity and briskness unattainable in his more pressured compositions. One of de Kooning's outstanding qualities (compared, for instance, to Pollock or Kline) is the variety of tempo he has been able to introduce into his action without destroying its continuity . . . His compositions devour everyday sights, odd thoughts, moods, theories old and new, paintings and sculptures of the past. He has the hungry multifariousness of the Renaissance humanists, the 'vulgarity' of Rabelais and Cervantes. (Pp. 104, f. Quoted in Alan Tormey, 'Critical Judgements', *Theoria,* 39, 1973, p. 48.)

The highly figurative character of much criticism is made the more startling for the absence from the serious critic's working vocabulary of such very general words as 'beautiful' and 'ugly'. People have long since come to remark the remoteness from actual criticism of discussions of the logic of 'good' and of 'beautiful' in aesthetics, and to realize what a severe handicap to criticism a vocabulary consisting of dichotomous pairs like 'good'/'bad' and 'beautiful'/'ugly' would be. Apparently a critic needs a much more refined and flexible set of tools than this, a rich and vital vocabulary, capable of supporting, so it seems, an indefinite number of critical distinctions. Perhaps it also seems that a vocabulary rich enough for the critic's purposes yet consisting of terms applicable without change in meaning to more than one aesthetic object is nowhere to be had.

Finally, there is a sense in which remarks such as those quoted above depend for their being understood on the reader's relatively sophisticated first-hand familiarity with the work or works under discussion. Thus Isenberg

> It seems that the critic's *meaning* is 'filled in', 'rounded out', or 'completed' by the act of perception, which is performed, not to judge the truth of his description, but in a certain sense to *understand* it (p. 163).

and

> Reading criticism, otherwise than in the presence, or with direct recollection, of the objects discussed is a blank and senseless employment (p. 164).

On these bases, Isenberg goes on to speculate that

> the meaning of [even] a word like 'assonance' . . . is in critical usage never twice the same (p. 165).

Isenberg is not here reporting the results of an exhaustive review of occurrences of the word 'assonance' in criticism. This is no mere factual claim to be overturned by a subsequent discovery or with the appearance of a new book by Northrop Frye. It presents a distinction between ordinary descriptive discourse and art criticism: in effect the NC thesis. As an explanatory hypothesis the NC thesis would do for the observations about art criticism mentioned above, if it weren't so counter-intuitive.

First of all, it just seems to be false that art-critical terms change meaning from employment to employment. Suppose we have two pieces of melancholy music. Does 'is melancholy' in 'Mozart's G Minor Quintet is melancholy' mean something different from what it means in 'Brahms' A Major Violin Sonata is melancholy'? We think not, but even if it does, the NC thesis requires still more. It requires that 'is melancholy' means something different in 'Brahms' A Major Violin Sonata is melancholy' from what it means in 'Brahms' D Minor Violin Sonata is melancholy'. And that is surely false.

Secondly, it is *prima facie* puzzling to talk of meaning in art criticism and of art-critical communication on the NC thesis. One wants to know how the critic can hope to communicate using words whose "meanings" cannot in principle be fixed. Perhaps this puzzle can be solved. However, we prefer to finesse it.

Let us distinguish between determinate and (sub-)determinable features as follows: A determinate feature is such that two or more objects which have it must be in that respect qualitatively indistinguishable. A (sub-)determinable feature is such that two or more objects which have it need not be in that respect qualitatively indistinguishable. Thus 'shaped' names a determinable feature, 'elliptical' a sub-determinable feature and 'square' a determinate feature; and a blue object will have the determinable feature of being coloured, the sub-determinable feature of being blue and a determinate feature shared only by things from which it cannot be distinguished by colour.[1]

Now, an object's determinate colour, for instance (even if the object can be distinguished by colour from everything else in the universe) is shareable. Thus, with colours, one *describes* an object whether one is talking about its sub-determinable *or* its determinate colour; and there are *both* sub-determinable and determinate colour concepts. But perhaps there is reason to suppose this not to be the case with art criticism. Perhaps a case can be made for the view that a work of art's determinate aesthetic features are not shareable. Clearly Isenberg believes that such a case can be made. But when he claims that the meaning of a word like 'assonance' is in critical usage never twice the same, he appears to move from the assumed non-shareability of determinate aesthetic features to the non-shareability of aesthetic features. But this does not follow. Nor does it follow from there being no determinate art-critical concepts that there are no art-critical concepts; there might still be sub-determinable art-critical concepts.

---

[1] The terms 'determinate' and 'determinable' originate in W. E. Johnson, *Logic* (Cambridge: Cambridge University Press, 1921), Part I, Ch. XI. We follow the usage of C. W. K. Mundle in his *Perception: Facts and Theories* (London: Oxford University Press, 1971), p. 133. The difference between determinable and sub-determinable features (both on the same side of the main distinction) may be construed in a number of ways, but it involves at least a difference in level of generality.

There is perhaps an explanation for this *non-sequitur*. If one held that the *point* of art criticism was specifically not to draw attention to similarities between works of art which share sub-determinable aesthetic features, but rather to draw attention to the determinate aesthetic features of each, and if one held in addition that the determinate aesthetic features of works of art were non-shareable, one might be inclined to re-interpret all critical reference to a work's sub-determinable aesthetic features as part of a necessarily round-about way of indicating what its non-shareable determinate aesthetic features are.[1] Still one would not be able to conclude that in art criticism the work of art is not, strictly speaking, described. For it might be objected that as part of the necessarily round-about way of indicating a work's non-shareable determinate aesthetic features, one finds oneself describing it in terms of its shareable sub-determinable ones.

---

[1] Cf. Isenberg's treatment of Goldschieder's critical remarks on *The Burial of Count Orgaz* at p. 162.

Dartmouth College and
the University of British Columbia

© JOEL RUDINOW and R. I. SIKORA 1975

DETERMINISM AND KNOWLEDGE

By William Lyons

(1) AN increasing number of philosophers have argued that genuine knowledge is impossible if determinism is true because genuine knowledge implies genuine truth which in turn implies freedom; the freedom to hold that a proposition is true for good reasons or the freedom to arrive at the correct belief about something or the freedom to make a claim that something is true or the like. Thus Warner Wick writes

> The upshot of my argument, and of these examples of "correctness", is that all talk of truth (or even of art) would be utterly *pointless* if there were nothing for it but causal influences that induced me to say or think *this* . . . ('Truth's Debt to Freedom', *Mind,* Vol. 73, 1964, p. 535).
>
> Correctness which is not self-determined is not truth (ibid., p. 537).

In his article 'The Conceivability of Mechanism' Norman Malcolm writes in similar vein

> . . . my acceptance of mechanism as true for myself would imply that I am incapable of saying or doing anything for a reason . . . it would also imply that I am incapable of having rational grounds for asserting anything including mechanism (*Philosophical Review*, Vol. 77, 1968, p. 70).

J. R. Lucas in his book *The Freedom of the Will* (Oxford: At the Clarendon Press, 1970) makes this point even more bluntly.

> Determinism, therefore, cannot be true, because if it was, we should not take the determinists' arguments as being really arguments, but as being only conditioned reflexes. Their statements should not be regarded as really claiming to be true, but only as seeking to cause us to respond in some way desired by them. . . . And therefore I cannot take determinism, whether in the mouth of another or believed by myself, seriously; for if it were true, it would destroy the possibility of its being rationally considered and recognised as such. Only a free agent can be a rational one. Reasoning, and hence truth, presupposes freedom just as much as deliberation and moral choice do (p. 115).

And there are other philosophers who argue in a similar way (for example, J. D. Mabbott, *An Introduction to Ethics,* London, Hutchinson, 1966, pp. 115–116; E. L. Mascall, *Christian Theology and Natural Science,* London, Longmans, pp. 212–219).

(2) In short, an increasing number of philosophers have argued that the determinist position meets its Waterloo in the area of knowledge, for the determinist is forced to hold two incompatible beliefs

(a) that our knowledge is determined

and

(b) that we have genuine knowledge.

Antagonists of determinism, such as Wick, Malcolm and Lucas, hold that if (a) is true then (b) is false, and that if (b) is true then (a) must be false. In pressing this dilemma upon the determinists these antagonists are themselves implicitly or explicitly espousing a proposition which underpins this dilemma but which also makes them open to counter-attack from the determinists. This proposition is

(c) that genuine knowledge is dependent upon truth which in turn presupposes freedom.

In this article I am going to argue that proposition (c) is false and that in consequence there is no incompatibility between propositions (a) and (b) above. If knowledge is partly defined in terms of the truth of a proposition, and if truth is the state of having so "hit the mark" that in the case of some descriptive proposition, for example, the proposition correctly describes what it claims to be a description of, then it should follow that truth is independent of *the way* in which "the mark is hit". But I will endeavour to argue this case in detail in the next section. I will also argue in a later section that the specific method of attaining truth which is called 'knowledge' is not incompatible with determinism.

(3) I can score a bull's eye on a target with an arrow by aiming carefully, holding the bow steady, and then letting the arrow fly with a smoothness bred of veteran skill. I can also score a bull's eye by falling over just as I am about to enter the archery contestants' circle in such a way that the arrow flies up into the air with a very high trajectory but somehow manages "by chance" to land in the bull's eye. A mechanical cross-bow, after being set upon the ground in the contestants' circle, pre-aimed, primed and connected to a timing device, can let fly an arrow and score a bull's eye.

I want to suggest that the last picture, that of the primed cross-bow hitting the target with an arrow, may well be a fair analogy of the determinists' view of knowledge and so help us to see that it is possible to have both knowledge and truth independently of freedom.

Let us first of all consider the truth step, for I take it that whatever is known is true but not vice versa. A chance falling together of the words of a word-game might form a true proposition and so be said to be true but it would not be said to be known. Knowledge includes other conditions centred on the way or manner in which the true proposition is arrived at.

If truth is a property of a proposition in the sense that it is a value we put on a proposition that has "hit the mark", or accurately arrived at something it claimed to arrive at, or corresponded to something, or the like, then it is strictly speaking irrelevant to cast doubt on its truth on account of the way it was arrived at. Truth arrived at in a causally determined way would seem to be just as true as truth arrived at in any other way. We do not seem to need to be free or self-determined in order merely to make the necessary moves to arrive at truth. Indeed there does not have to be a "we" around. Truth can be arrived at independently of persons, though, as we shall see, I do not think knowledge can. A computer, the "chance" falling together of the words of a word-game or a parrot can produce true propositions or truth.

I suggest that what the philosophers I actually quoted, particularly Wick and Lucas, ought to be claiming is not that truth presupposes freedom, but that knowledge does, for I think that this way of putting the case at least seems plausible. If freedom enters into knowledge at all—and I will argue it need not—it will make sense only at the method level, not at the end-product or truth level. It will make sense at the method level because freedom is opposed to determinism, and determinism is a thesis about things resulting causally from other things, that is, determinism is a thesis about the way or method in which things come about. Likewise freedom must be a thesis about the way in which things come about.

(4) So now I want to make out a case that, while knowledge is not independent of the method of arriving at a true proposition, it is independent of the nature of the original impetus of the method. Ultimately, the dispute between determinists and self-determinists in the context of knowledge will be a dispute about whether the particular way of arriving at a true proposition which categorizes the process as one of knowledge will be altered if one has as the instigator of these steps causal determinism rather than free choice. I want to argue that it makes no difference: an advocate of the view that knowledge involves freedom should argue that it does.

Besides the stipulation that a *true proposition* has been arrived at or produced, knowledge must include some such conditions as that a (human?) person has arrived at the true proposition by way of responding to reasons or evidence or the like, and that the person is aware that he has arrived at the true proposition on account of the reasons, evidence or the like. I think that Lucas, Malcolm and Wick are correct in thinking that the vital condition that transforms truth into knowledge is the dependence of knowledge upon the method of arriving at truth by way of *reasons*. I would add that the notion of knowledge is also tied up with the *realization* by the possessor of the reasons that they are good or *justifying* in some sense. I suspect that knowledge may well be tied to some notion of *person*, and perhaps to one of *human* person as well, but I do not think that these are the crucial considerations for the dispute under discussion.

I do not see that we need to be free in order to be conscious of or realize that we have arrived at the truth of some proposition for good reasons or the like. There seems to be no logical objection to the possibility that we are so determined that when presented with, say, a descriptive proposition we look for reasons to justify our holding or not holding it as a correct description of some aspect of the universe and, when we have found such reasons and realize that we have found them, we claim knowledge and realize that we are doing so. We may be mistaken, of course, because our mechanism has gone wrong in some way. We may, so to speak, have "gone off half-cocked" and claimed knowledge without sufficient reason, or for no reason, or for an irrelevant reason, and so on. In short, there seems to be no logical objection to the possibility that a determinist can be caused to be conscious of having arrived at a proposition which is true by a rational process. It may be the case that we are highly sophisticated mechanisms primed to sort out, adjudicate about the merits of, decide on the relevancy of, and finally respond causally to, certain sorts of reasons for holding propositions.

I do not pretend to have given more than a few necessary conditions for knowledge, in fact by and large I have taken the word of Lucas, Malcolm and Wick that the important condition for knowledge is that it is a way of arriving at truth rationally. Besides, the important point is that there does not seem to be any good reason to believe that, whatever might be the look of a list of the necessary and sufficient steps which make up the method for arriving at truth in such a way that the whole process is called 'knowledge', one could not simply add in front of the list of the steps, 'the person was causally determined to . . .'. The list of steps should be independent of the nature of the authorship or origin of the steps, and so the listing of the steps themselves should not be able to rule out the possibility that their origin is causally determined.

(5) Now I want briefly to discuss what are often taken to be alternative versions of the Wick, Malcolm, Lucas position.

It is sometimes suggested that to be causally determined to hold that knowledge is determined is to be in such a position that one would never know whether what one holds is true or not, and so would never know whether determinism is true or not. But this it seems to me is to mistake the nature of causality.

This thesis seems to imply that causality is inimical to awareness, in such a way that if x is caused to do y then it follows that x cannot realize or be aware that he is doing y. This view may well stem from a belief that all causality must be at the billiard-ball level, so that to say x caused y is immediately to limit the filling of the variable gaps x and y to unconscious insensible items and items shorn of their conscious powers. But why should this be so? Why can't x be an item possessing consciousness or awareness, so that when x is caused to do y, the possibility is left open that x also be caused to be aware of doing y? If items possessing consciousness or awareness are items that can be isolated and individuated in space-time, then why should they not be amenable to causality in respect of all their powers?

(6) Finally, it is no objection against determinism to suggest that, if we are determined we might be so "wound up" by evolution or our genetic make-up or the like that we continually (or very often, or now and again) believe to be true propositions which may be false, though we would never know. For this objection (a version of the Cartesian 'evil demon') is also an objection against self-determinism. One could equally say that, if we are self-determined, we might continually be causing ourselves to believe to be true propositions which may be false, though we would never know it. The fallacy about making this sort of move against determinists or anyone else is that, if we don't *now* possess the apparatus for unequivocally deciding what is true and false, who does? Where do our notions of unequivocal truth and falsity come from? What sense could be given to 'truth' unless it is something that we *now* unequivocally arrive at and *now* unequivocally distinguish from falsity? Wouldn't it mean that we would have to hold a distinction between 'truth' and 'real truth'? Descartes, of course, had God in the background as the final arbiter of what is "really" true and false, and so his radical doubt about the possibility of a deceitful evil demon was made initially plausible. In this agnostic age, this is not open to us. Besides, most theists would agree that even if God is about and does possess the absolute yardstick of what is true and false, He doesn't tell us!

Most importantly, those who press the above objection against determinism fail to see that it is not so much an objection against determinism as an objection against any theory of truth and knowledge. As Berkeley said, an objection against any theory at all is not an objection against any particular one.

University of Glasgow

© William Lyons 1975

HART ON PATERNALISM

By Christine Pierce

IN *Law, Liberty, and Morality* (New York: Vintage Books, 1963) H. L. A. Hart defends J. S. Mill's liberal view of liberty against Patrick Devlin's claim that there can be no theoretical limits to the power of the State to legislate against immorality. It is clear that Mill's principle of liberty was designed to function as just such a limit; however, it is unclear how liberal Hart intends to be, for he outlines at least two positions on paternalism which have conflicting and, in some cases, peculiar consequences.

Mill held that if an individual's actions harm no other party, those actions should not be considered criminal even though they are viewed by many as unwise, disgusting, ridiculous or even immoral. Hence Mill's liberalism results in anti-paternalism: provided that others are not harmed, self-regarding acts and acts consented to, even if considered harmful to oneself or those who consent, are matters of privacy or personal liberty.

Hart characterizes Mill's protests against paternalism as 'fantastic', and suggests a modified principle of liberty which accommodates paternalism by protecting consenting victims without condoning the legal moralism of Devlin. According to Devlin, consent should not be a defence against euthanasia because society has the right to enforce its morals legally. For Hart, the reason a decision regarding euthanasia cannot be left to the individual is not that society may view the practice as immoral. Hart's reason is that people may harm themselves, and his concern that wrong decisions will be made is largely based on his distrust of consent. He suggests that if we no longer sympathize with Mill's anti-paternalism this is partly due to '. . . an increased awareness of a great range of factors which diminish the significance to be attached to . . . consent' (pp. 32, sq.). For example, consent may be given '. . . without adequate reflection or appreciation of the consequences; or in pursuit of merely transitory desires; or in various predicaments when the judgment is likely to be clouded . . .' (p. 33). Concerning laws against euthanasia Hart says, 'This is a rule of criminal law which many now wish to retain, though they would also wish to object to the legal punishment of offences against positive morality which harm no one' (p. 30). Again Hart differentiates his view from Devlin's by claiming that statutes punishing cruelty to animals are concerned with the '. . . suffering, albeit only of animals, rather than the immorality of torturing them' (p. 34). Certainly, he continues, 'no one who supports this use of the criminal law is thereby bound in consistency to admit that the law may punish forms of immorality which involve no suffering to any sentient being' (ibid.).

Hart's position, then, is that the criminal law may not be used to enforce morality as such. However, paternalism is justified insofar as it prevents people from making decisions which may result in their own harm. When restating this view later on, he substitutes 'suffering' for 'harm'—a harmless enough move it might seem, but it has significant consequences. Hart does not indicate whether or not he thinks that 'harm' would differ from 'suffering' and if so how.

I can imagine being told that I ought not to read pornography because it will harm me. It is difficult to argue with people who think your soul or your personality is being harmed. Saying you do not think so never seems to be enough. But if their claim is that it will make me suffer, and I say it will not, it is not so easy for them to think, 'She is suffering even though she believes she is not'. To cite another example, it is easier to defend homosexuality against the charge of suffering than against the charge of harming oneself. It is clear that people of the same sex can love each other and not be suffering.

Traditionally paternalists have argued against euthanasia, presumably on the grounds that nothing could be worse for you, i.e., more harmful, than dying. Hart, as we have seen, finds this notion plausible. However, according to the view that the criminal law should be used to prevent suffering, euthanasia is clearly justified. In espousing the 'harm' criterion for a paternalistic policy, Hart recognizes the reasonableness of retaining laws against euthanasia. At the same time, he is sketching another view which justifies euthanasia, failing to take account of the fact that a choice between the 'harm' criterion and the 'suffering' criterion could mean the difference between condemning or favouring euthanasia. Furthermore, if the prevention of suffering is to take precedence over the principle of consent in the criminal law, a case could be made for making euthanasia compulsory.

Hart might escape this radical conclusion by claiming that his view is somewhat more complicated than disallowing consent to suffering. In introducing the idea that the harm or suffering of consenting victims must be caused by another person, it is unclear which of four possible positions he may be espousing

(1a) Harming others, even if they consent, is legally culpable

(1b) Consenting to being harmed is legally culpable

(2a) Causing others to suffer, even if they consent, is legally culpable

(2b) Consenting to being made to suffer is legally culpable.

Hart's paternalistic views do not constitute a full-blown doctrine; primarily he is responding to Devlin by suggesting that harm and suffering, not immorality, need to be prevented. When expressing this view, Hart could be interpreted as adopting 1b and/or 2b. However, on one occasion, Hart gives a more precise statement of principle which argues for 1a and/or 2a: '. . . harming others is something we may still seek to prevent by use of the criminal law, even when the victims consent to or assist in the acts which are harmful to them' (p. 33). Given Hart's willingness to substitute 'suffering' for 'harm' this quotation is evidence for 2a as well as 1a.

If Hart accepts 2a rather than 2b as an adequate statement of his 'suffering' criterion, he can escape the dramatic consequence of compulsory euthanasia for he would not be saying simply that a person may not choose to suffer, but that a person may not cause the suffering to which an individual consents. This principle would be largely irrelevant in deciding cases of euthanasia since the suffering involved is not normally caused by another human being. Nonetheless, no significant advance has been made toward a coherent theory of legal paternalism. Although 2a extricates Hart from ludicrous conclusions about euthanasia, it is nevertheless an unsatisfactory statement of principle for it produces equally absurd results in other contexts. For example, if a person dying a painless death could be restored to health by a cure that causes great suffering, no one would be allowed to effect that cure under 2a.

Whether or not a causality rider is included in Hart's paternalism, the major difficulty with his position is a failure to sharpen up the differences between harm and suffering, thus overlooking the unique problems of each criterion as well as the possibility that they might have conflicting consequences. It seems unfair to use 'harm' as a criterion for infringing on personal liberty, because the vagueness of it makes it difficult, if not impossible, to defend oneself. 'Suffering' is an attractive criterion and/or interpretation of 'harm' because it is easier to identify, but it has incredible consequences.

State University of New York, Oswego © CHRISTINE PIERCE 1975

LOCKE'S UNNATURAL KINDS

By Douglas Odegard

J. L. MACKIE rightly points out that Locke sees a connection between the practice of using substance names to refer directly to real essences and the belief in natural kinds ('Locke's anticipation of Kripke', Analysis 34.6, June 1974). Mackie endorses the latter belief, *as the belief that nature works regularly,* and includes it as part of his rationale for commending such a practice, contrary to much of what Locke holds. The question of whether the practice which Mackie commends is the same as the one Locke rejects is too complicated to resolve in a short space. But it is worth noting that the belief in natural kinds which Locke has in mind is not merely the belief in natural regularities. The relevant part of the *Essay* which Mackie quotes is

> That which I think very much disposes men to substitute their names for the real essences of species, is the supposition before-mentioned, that nature works regularly in the production of things, and sets the boundaries to each of these species, by giving exactly the same real internal constitution to each individual, which we rank under one general name. Whereas any one who observes their different qualities, can hardly doubt, that many of the individuals, called by the same name, are, in their internal constitution, as different from one another as several of those which are ranked under different specific names. (3.10.20.)

The 'before-mentioned supposition' to which Locke refers is one he discusses at length in Chapter 6 of Book 3, 'Of the Names of Substances'. The supposition has a number of elements, including the following.

(1) Two specimens of the same species cannot differ as much as two specimens of different species (3.6.8; also note the last sentence in the quoted passage).

(2) Different species are so distinct that cross-breeding is impossible (3.6.23).

(3) No two species are such that there can be an intervening "hybrid" species. Nor can there be monsters or changelings (3.6.12, 16, 17, 22, 23, 26, 27; also 3.3.17).

(4) No individual can be a member of more than one species (3.6.13).

We need not, and should not, accept *this* doctrine of natural kinds as partial justification for annexing substance names to real essences.

University of Guelph

© Douglas Odegard 1975

26 DEC